



Il ruolo della statistica nelle scienze dell'atmosfera: dall'osservazione all'interpretazione.

Piera Ielpo

p.ielpo@isac.cnr.it

con il contributo di Daniela Cava, Daniela Cesari, Rita Cesari,
Francesco Marra e Marco Paglione

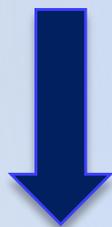
Sommario

- **Receptor oriented models** per la stima delle sorgenti di inquinanti e dei loro contributi (**source apportionment**) :
 - Receptor oriented models basati sul bilancio di massa
 - Receptor oriented models basati sull'analisi delle traiettorie delle masse d'aria.
- **Trasformate di Wavelet** applicate allo studio della turbolenza nel PBL
- **Metodi non asintotici** per lo studio delle precipitazioni estreme

- Receptor oriented models basati sul bilancio di massa al recettore

Source apportionment:

- Receptor oriented models basati sul bilancio di massa



Conservazione della massa tra la sorgente emissiva ed il sito

Input:

Composizione chimica di campioni di aerosol ambientali

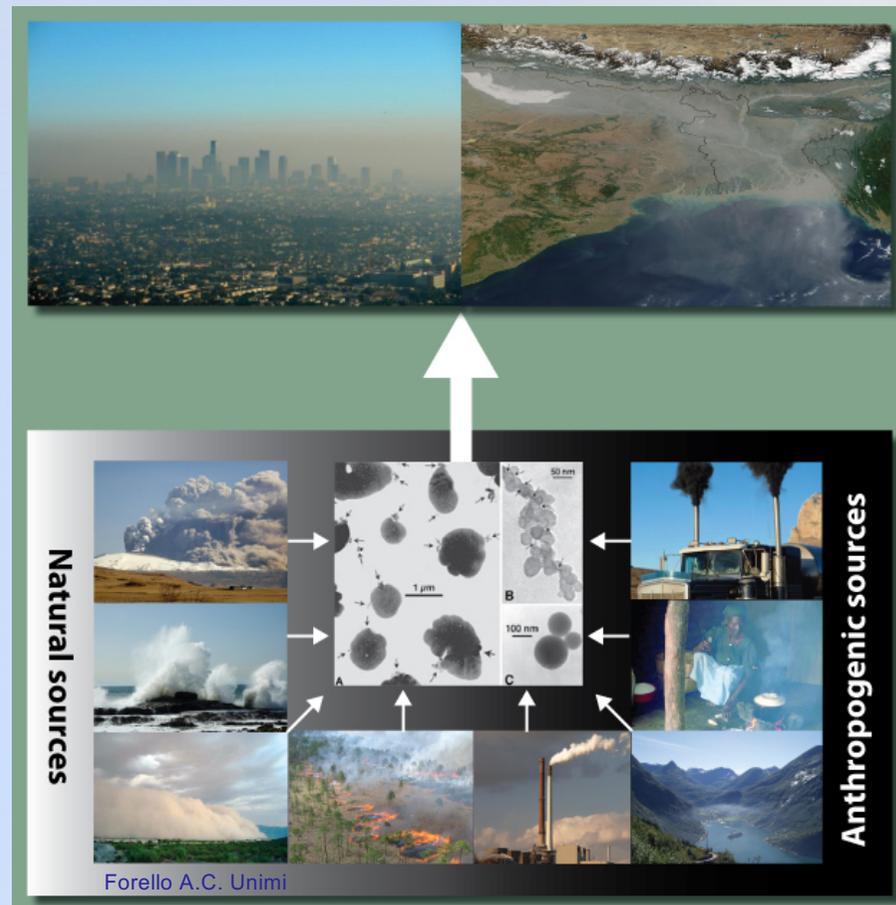


Source apportionment



Output:
Source temporal patterns

Source chemical profiles



Output:

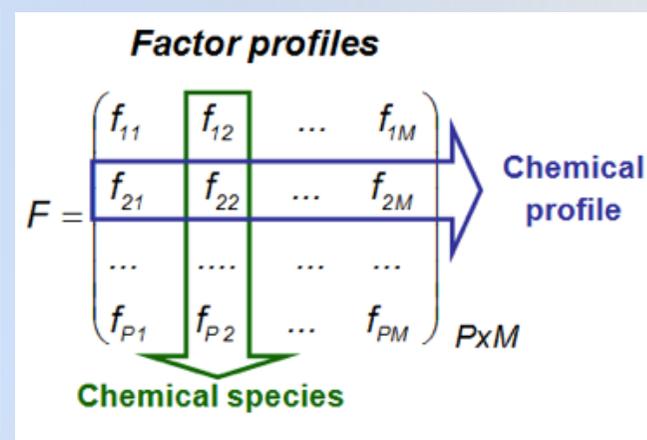
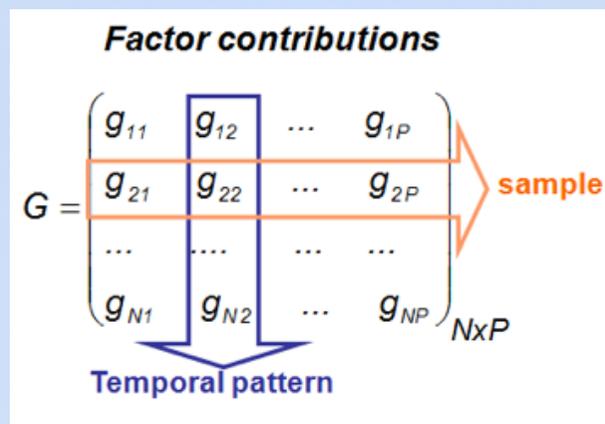
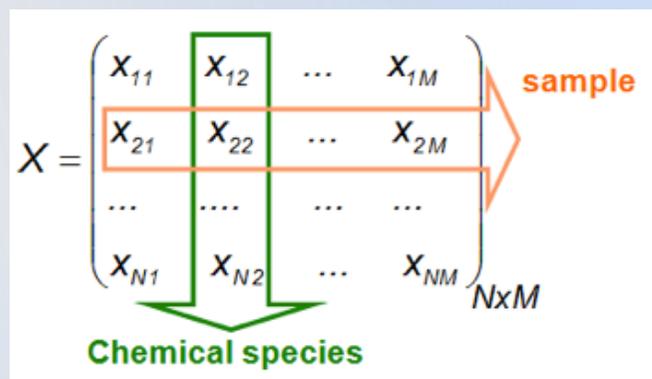
- Receptor oriented models basati sul bilancio di massa

Principi del receptor modelling

$$x_{ij} = \sum_{k=1}^P g_{ik} f_{kj} + e_{ij}$$

Equazione del bilancio di massa

$$X = G \cdot F + E$$



Fattori



Sorgenti

Interpretazione fisica

GOAL: Stimare i contributi delle sorgenti al sito recettore

Applicazione di PCA ed APCS: caso studio sulla città di Taranto

- Receptor oriented models basati sul bilancio di massa

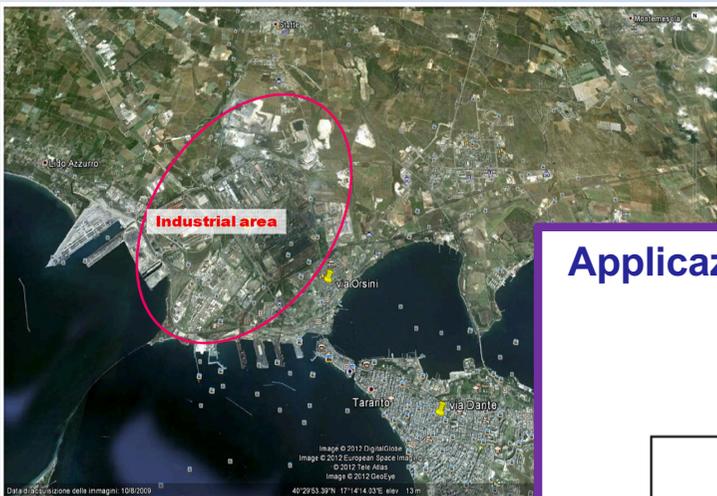
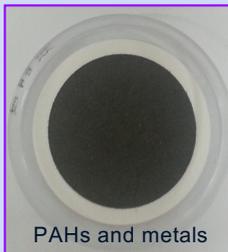


Fig. 1. Map of two sampling sites in Taranto: industrial site in Via Orsini and urban site in Via Galatina.



How a Steel Plant Affects Air Quality of a Nearby Urban Area and PAH Concentrations. M. Amodio...P. Ielpo et al. Aerosol and Air Quality Research, 13: 497–508, 2013

Applicazione del modello APCS a dati di caratterizzazione chimica di campioni di PM2.5 raccolti nel sito di via Orsini

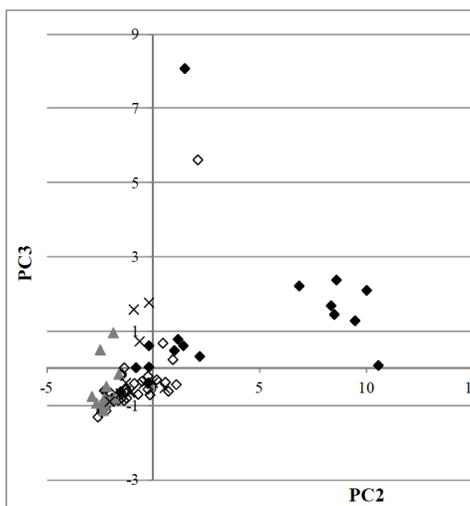
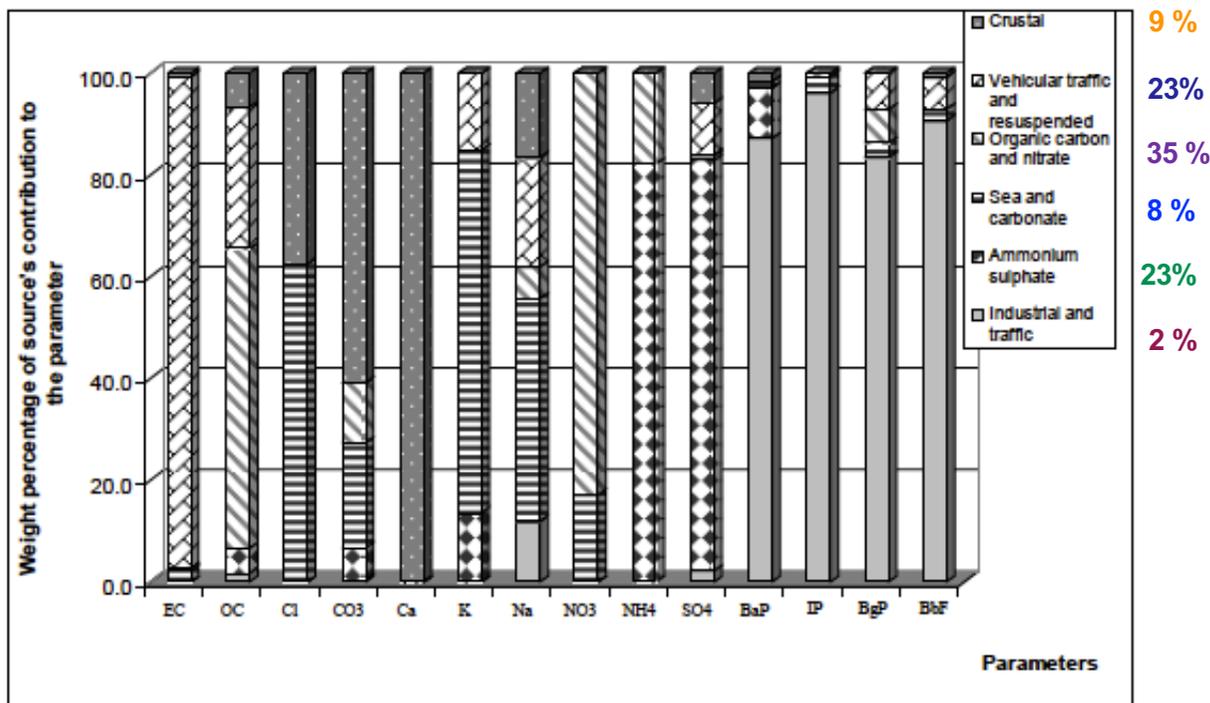


Fig. 7. PCA application to data (PAHs, Ca, Fe, Si, Zn, Mn) collected in the industrial site of Galatina and Torchiarolo (PC2 vs PC3).

M. Amodio...P. Ielpo et al. Chemical Engineering Transactions 2008

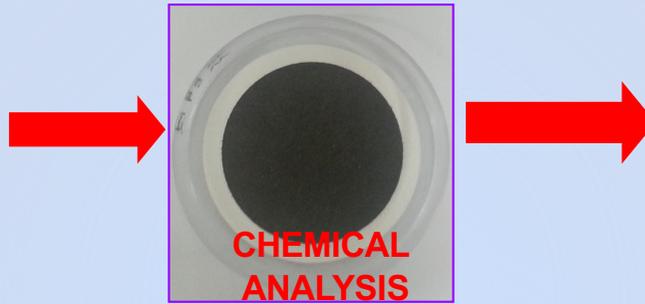
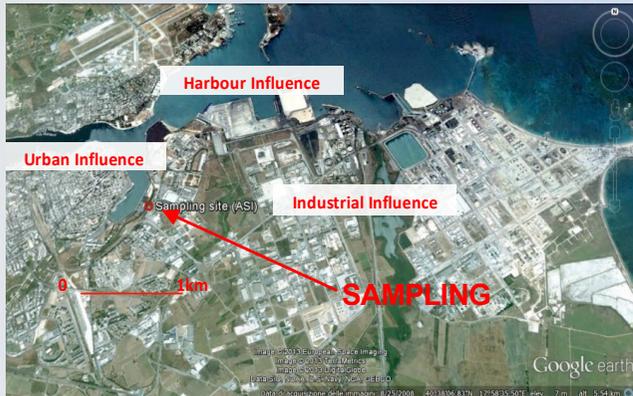


Positive Matrix Factorization model

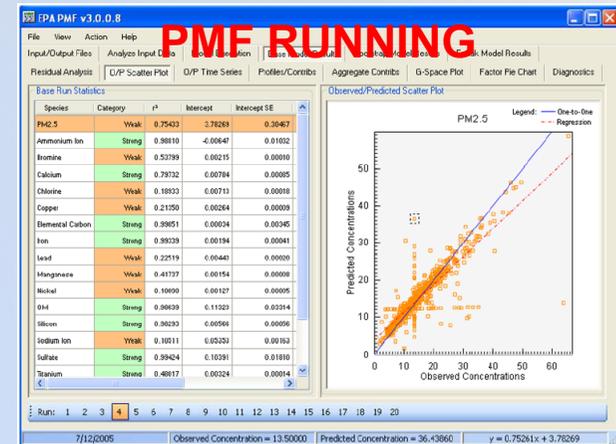
Caso studio: SA del PM_{2.5} nell'area portuale-industriale di Brindisi

PM_{2.5} medio 15.1 µg/m³.

• Receptor oriented models basati sul bilancio di massa



Massa PM_{2.5} caratterizzata: 51.4%



IL MODELLO A RECETTORE PMF FORNISCE UNA SOLUZIONE A 8 FATTORI/SORGENTI

	Source contribution (ng/m ³)	Source contribution (%)
Industrial/oil combustion	2300 (± 195)	15.3 (± 1.3)
Crustal	2475 (± 134)	16.4 (± 0.9)
Biomass combustion/fires	1772 (± 101)	11.7 (± 0.7)
Traffic	2474 (± 259)	16.4 (± 1.7)
Industrial	56 (± 48)	0.4 (± 0.3)
Sea spray	387 (± 77)	2.6 (± 0.5)
Secondary Ammonium Sulphate	4115 (± 120)	27.3 (± 0.8)
Crustal carbonates	1166 (± 42)	7.7 (± 0.3)

COME STIMARE IL CONTRIBUTO PRIMARIO DEL TRAFFICO NAVALE???

$$PM_a = \langle r \rangle * V_a / \langle F_{V, HFO} \rangle$$

- **PM_a** : PM_{2.5} primario (µg/m³);
- **r**: average ratio of PM_{2.5} to normalised V emitted according to the HFO burning experiment engines;
- **V_a**: concentrazione atmosferica di V (µg/m³);
- **F_{V, HFO}**: concentrazione media di V contenuta nel HFO usato dalle navi nei porti;

CONTRIBUTO PRIMARIO DEL TRAFFICO NAVALE AL PM_{2.5}: 2.8 %



Source apportionment of PM_{2.5} in the harbour–industrial area of Brindisi (Italy): Identification and estimation of the contribution of in-port ship emissions

D. Cesari ^{a,*}, A. Genga ^b, P. Ielpo ^{a,c}, M. Siciliano ^b, G. Mascolo ^c, F.M. Grasso ^a, D. Contini ^a

^a Istituto di Scienze dell'Atmosfera e del Clima, ISAC-CNR, 73100 Lecce, Italy
^b Dipartimento di Scienze e Tecnologie Biologiche e Ambientali, Università del Salento, 73100 Lecce, Italy
^c Istituto di Ricerca Sulle Acque, IRSA-CNR, 70132 Bari, Italy

d.cesari@isac.cnr.it

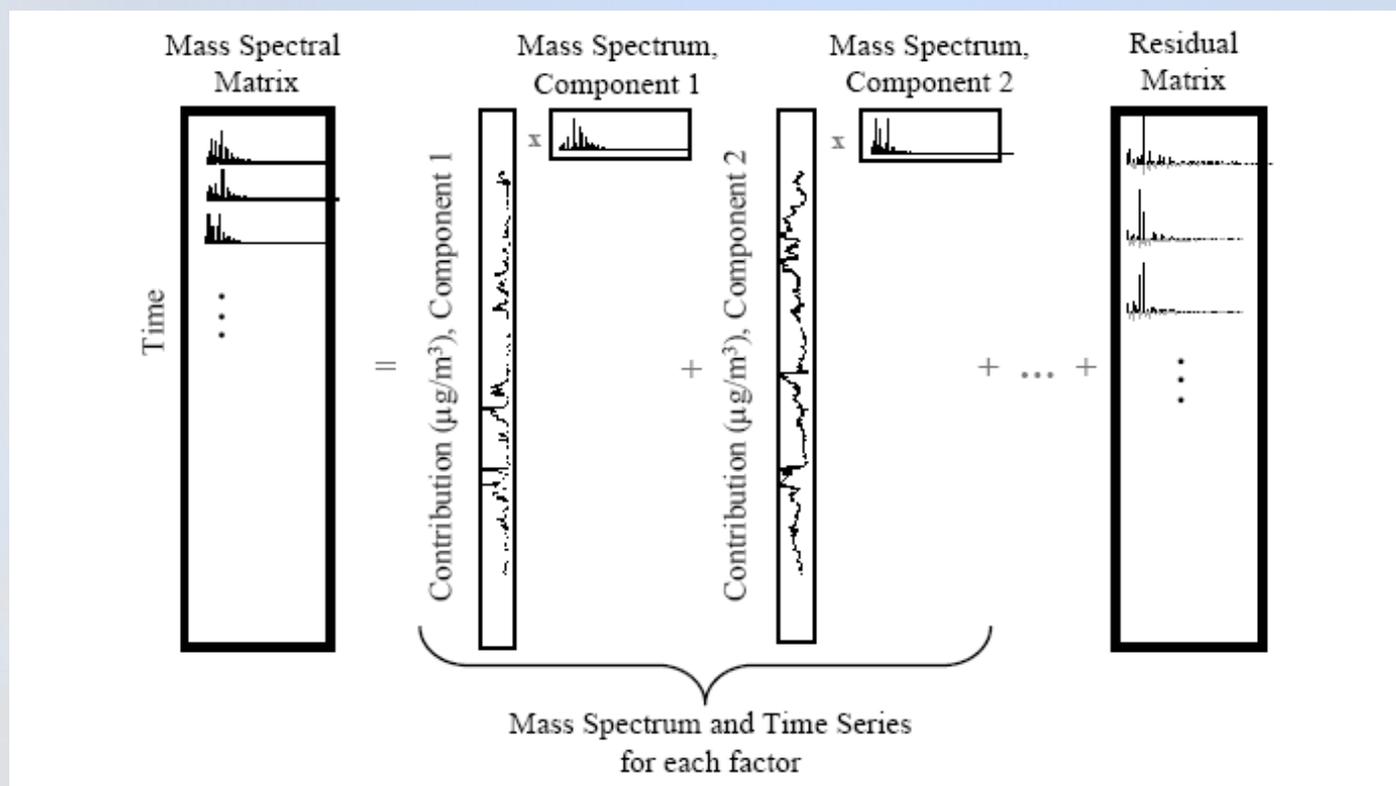
- Receptor oriented models basati sul bilancio di massa

Source apportionment applicato agli spettri di massa dell' aerosol organico

m.paglione@isac.cnr.it

Modelli a recettore basati su tecniche di statistica multivariata degli spettri di massa dell'OA:
Positive Matrix Factorization (PMF)

Serie di spettri di massa AMS dell'OA



Analisi fattoriale

$$(X = G * F + E)$$

Where:
 X = measurements matrix
 F = factor profile vector
 G = factor contribution vector
 E = errors matrix

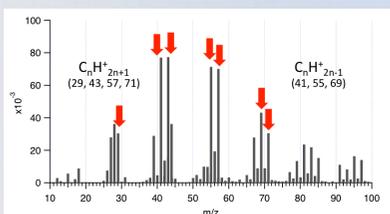
Principali algoritmi di calcolo utilizzati:

PMF2 & ME-2

- Receptor oriented models basati sul bilancio di massa

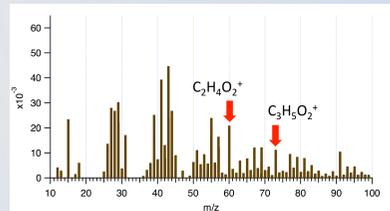
Principali sorgenti di aerosol organico (OA) identificate tramite AMS-PMF

Componenti Primarie (POA)



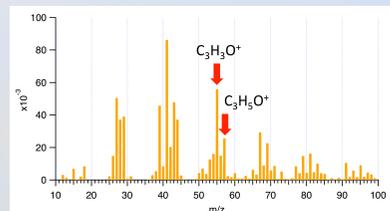
HOA
(Hydrocarbon-like)

Combustione combustibili fossili = TRAFFICO



BBOA
(Biomass-burning)

Combustione di biomassa = RISCALDAMENTO DOMESTICO



COA
(Cooking)

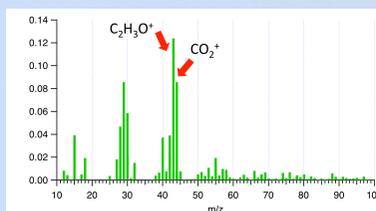
Cottura dei cibi = CUCINA

Componenti Secondarie

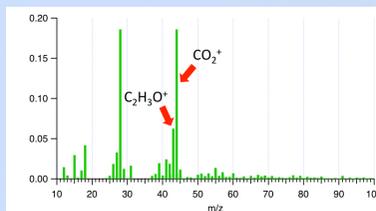
i cosiddetti Oxidized Organic Aerosol (OOA), proxy dei

SOA

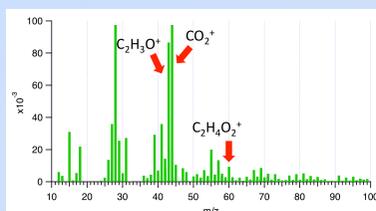
SV-OOA
(semivolatile Oxidized)



LV-OOA
(lowvolatile Oxidized)



OOA-BB
(biomass-burning Oxidized)



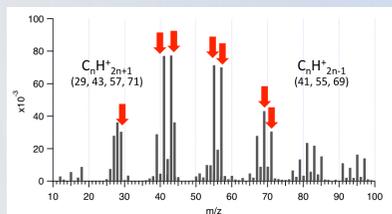
Molteplicità di processi di formazione e di trasformazione

...

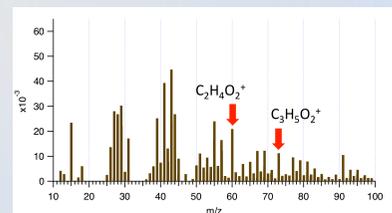
- Receptor oriented models basati sul bilancio di massa

Principali sorgenti di OA identificate tramite AMS-PMF

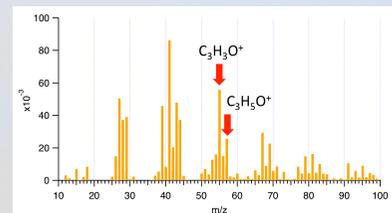
Componenti Primarie (POA)



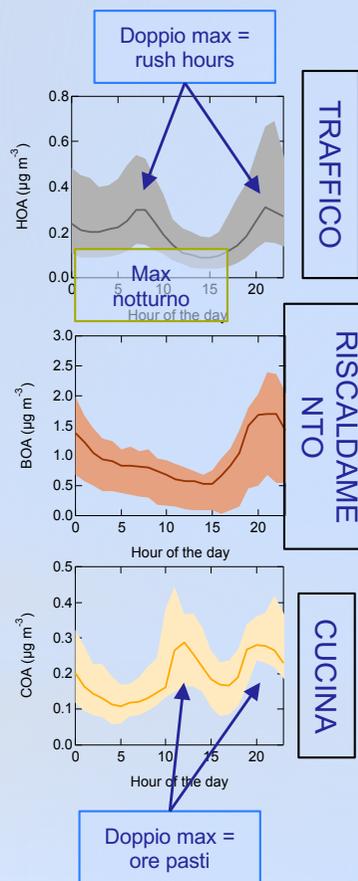
HOA
(Hydrocarbon-like)



BBOA
(Biomass-burning)

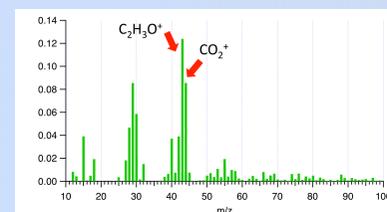


COA
(Cooking)

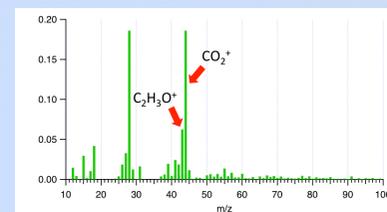
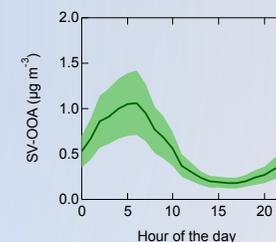


Componenti Secondarie

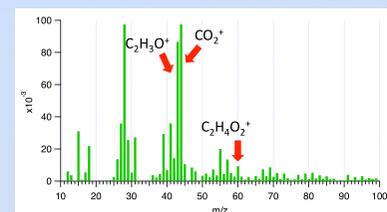
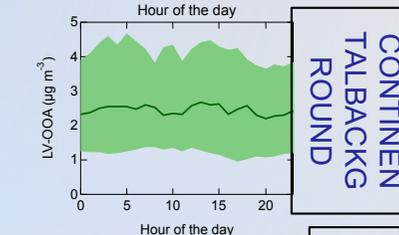
i cosiddetti Oxidized Organic Aerosol (OOA), proxy dei SOA



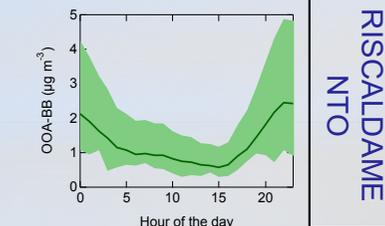
SV-OOA
(semivolatile oxidized)



LV-OOA
(low-volatility oxidized)



OOA-BB
(biomass-burning oxidized)



- Receptor oriented models basati sul bilancio di massa

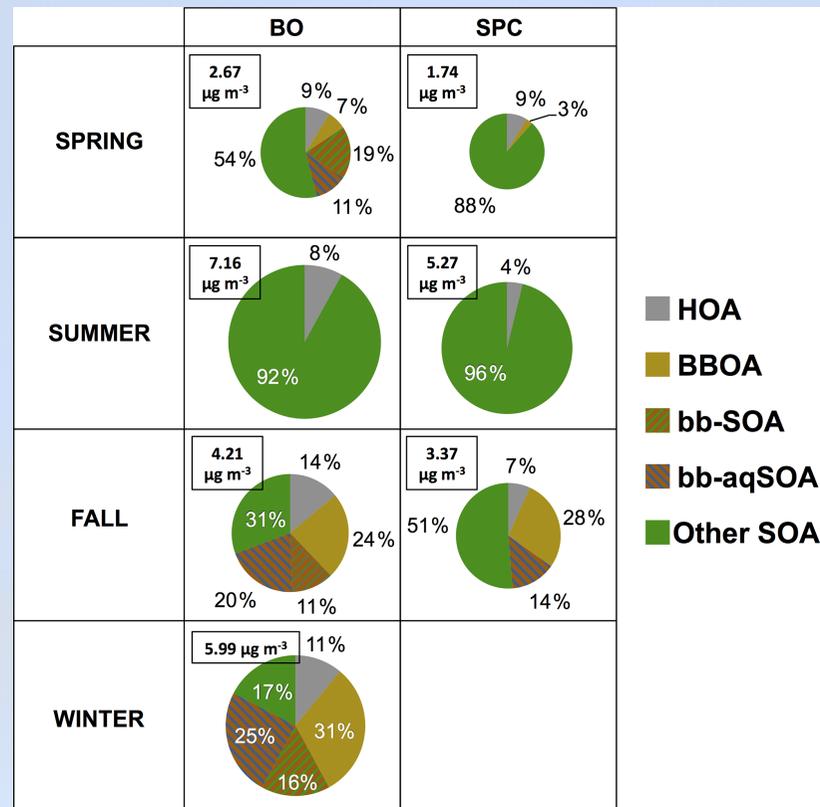
Progetto Supersito: interazione tra qualità dell'aria e salute umana



Contributi delle principali sorgenti di aerosol organico in Emilia Romagna (2011-2014)

Enti pubblici e di monitoraggio ambientale e istituzioni di ricerca nazionali ed internazionali collaborano per migliorare le conoscenze relative agli aspetti ambientali e sanitari del particolato fine (PM_{2.5} e PM₁) e ultrafine in atmosfera e negli ambienti chiusi.

Paglione et al., ACP, 2020



Nelle **STAGIONI FREDE**: contributo importante della frazione primaria originata dalla **combustione di biomassa**, che influenza molto anche la **FRAZIONE SECONDARIA (SOA)**.

In **ESTATE**: La combustione di biomassa scompare e diventa preponderante (quasi totale) la **frazione secondaria => processi fotochimici**

Some more references:
 Gilardoni et al., ACP, 2014;
 Gilardoni et al., PNAS, 2016;
 Costabile et al., ACP, 2017;
 Paglione et al., ACP, 2020;

...



Source apportionment: Receptor oriented models basati sull'analisi delle traiettorie delle masse d'aria

Analisi statistica delle traiettorie all'indietro (Trajectory statistical methods (TSMs)).
Modello a traiettoria utilizzato: FLEXPART (Stohl et al., 2005, Pisso et al., 2019)

- **Residence Time (RT) Analysis** (Ashbaugh L.L., 1983):

$$P[A_{ij}] \approx \frac{n_{ij}}{N}$$

dove n_{ij} è il numero di posizioni delle particelle d'aria che, ad ogni incremento temporale, cadono nella ij^{th} cella durante un intervallo di tempo T , ed N è il numero totale di punti finali calcolati durante l'intervallo di tempo.

$P[A_{ij}]$ rappresenta quindi il **tempo di residenza (residence time)** di una particella d'aria, selezionata in modo casuale, nella cella ij^{th} relativamente all'intervallo totale di tempo T .

- **Potential Source Contribution Function (PSCF)** (Zeng and Hopke, 1989):

$$P[B_{ij} | A_{ij}] = \frac{P[B_{ij}]}{P[A_{ij}]} = \frac{m_{ij}}{n_{ij}} = \frac{\sum_{l=1}^{M_l} \tau_{ijl}}{\sum_{l=1}^M \tau_{ijl}}$$

dove n_{ij} è il numero di posizioni finali nella ij^{th} cella durante il periodo T , m_{ij} sono le posizioni finali delle traiettorie che partono dal recettore quando la concentrazione di inquinante è maggiore del valore assegnato, M è il numero totale di traiettorie, M_l è il numero totale di traiettorie che arrivano al recettore quando la concentrazione è superiore al valore di soglia, e τ_{ijl} il tempo speso nella cella (i,j) dalla traiettoria l .

- Receptor oriented models basati sull'analisi delle traiettorie delle masse d'aria

- **Concentration Field method (CF)** (Siebert, P. et al., 1994):

$$\overline{\log C_{ij}} = \frac{1}{\sum_{l=1}^M \tau_{ijl}} \sum_{l=1}^M \log(c_l) \tau_{ijl}$$

- **Peak Events Concentration Field (PECF)** (Cesari et al., 2013):

$$\overline{\log C_{ij}} = \frac{1}{\sum_{l=1}^{M_t} \tau_{ijl}} \sum_{l=1}^{M_t} \log(c_l) \tau_{ijl}$$

r.cesari@isac.cnr.it

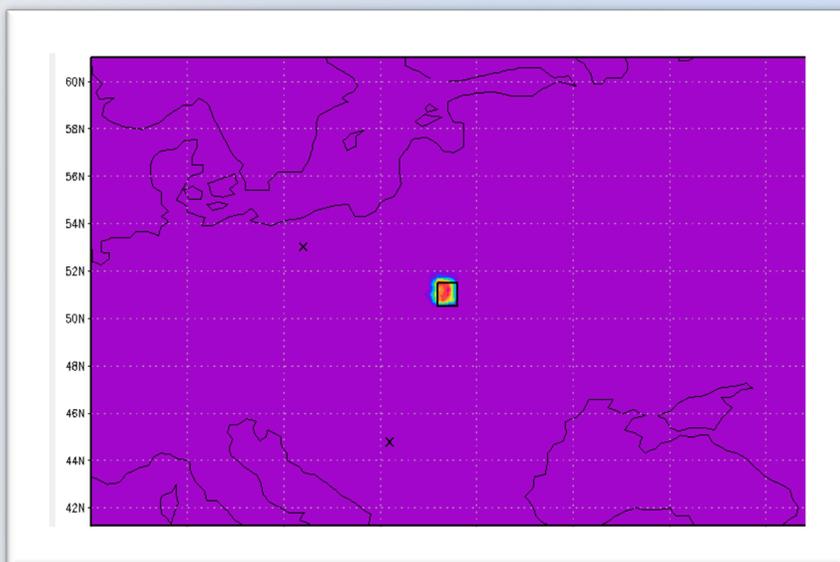
dove l è l'indice della traiettoria, M il numero totale di traiettorie, M_t il numero totale di traiettorie che arrivano al sito recettore quando la concentrazione dell'inquinante è maggiore di un assegnato valore di soglia, c_l la concentrazione misurata all'arrivo della traiettoria l e τ_{ijl} il tempo speso nella cella (i,j) dalla traiettoria l .

Siebert, P. et al., 1994, in Borrell, P.M., Borrell, P., Cvitai, T., Seiler, W. (Eds.),

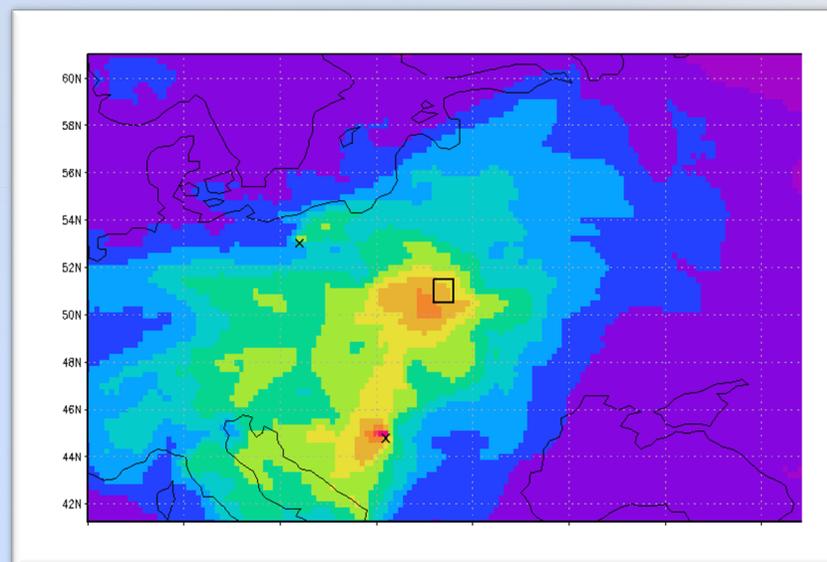
Transport and Transformation of Pollutants in the Troposphere. Academic Publishing, Den Haag, 689–693

Cesari et al., 2014, *Int. J. Environment and Pollution*, Vol. 55, Nos. 1/2/3/4

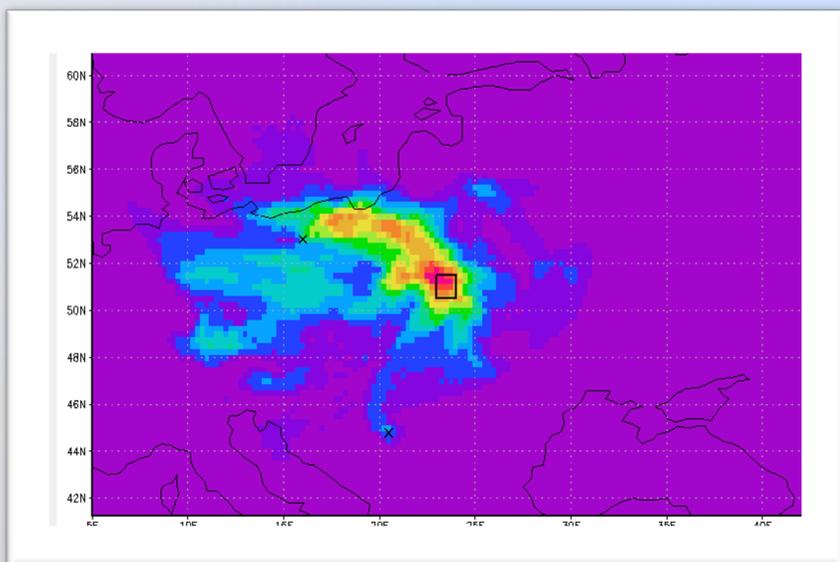
- Receptor oriented models basati sull'analisi delle traiettorie delle masse d'aria



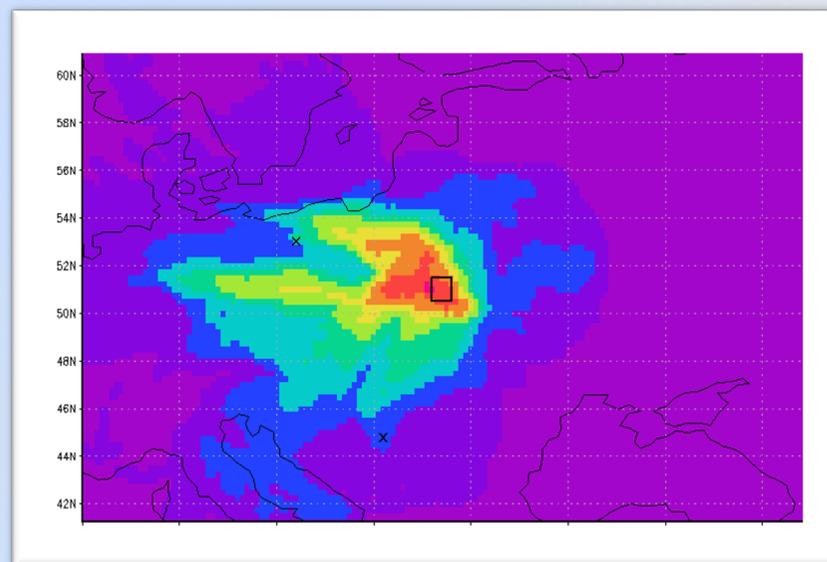
(a)



(b)



(c)



(d)

Dominio di simulazione. Il quadrato indica la sorgente, la croce indica i recettori (a). Ricostruzione della distribuzione spaziale della sorgenti con il metodo **CF** (b), **PSCF** (c) e **PECF** (d). (R. Cesari et al., 2013)

- Receptor oriented models basati sull'analisi delle traiettorie delle masse d'aria

Metodo bayesiano di inversione per la stima di flussi di inquinanti.

(collaborazione con UNIURB (M. Maione e U. Giostra) e Università La Sapienza)

Modello utilizzato: FLEXINVERT+ (Thompson and Stohl, 2014)

Per stimare i flussi di inquinanti, si utilizza il modello FLEXINVERT+.

Consideriamo:

$$\mathbf{y}^{mod} = \mathbf{H}\mathbf{x}$$

FLEXINVERT+ , avendo in input:

- mappe di sensitività (SRRs);
- mappe di emissioni (campo a-priori);
- concentrazioni misurate ai punti recettori;
- concentrazione di Background;
- relative incertezze;

minimizza quindi la funzione costo definita dall'equazione:

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} (\mathbf{H}\mathbf{x} - \mathbf{y}^{obs})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y}^{obs})$$

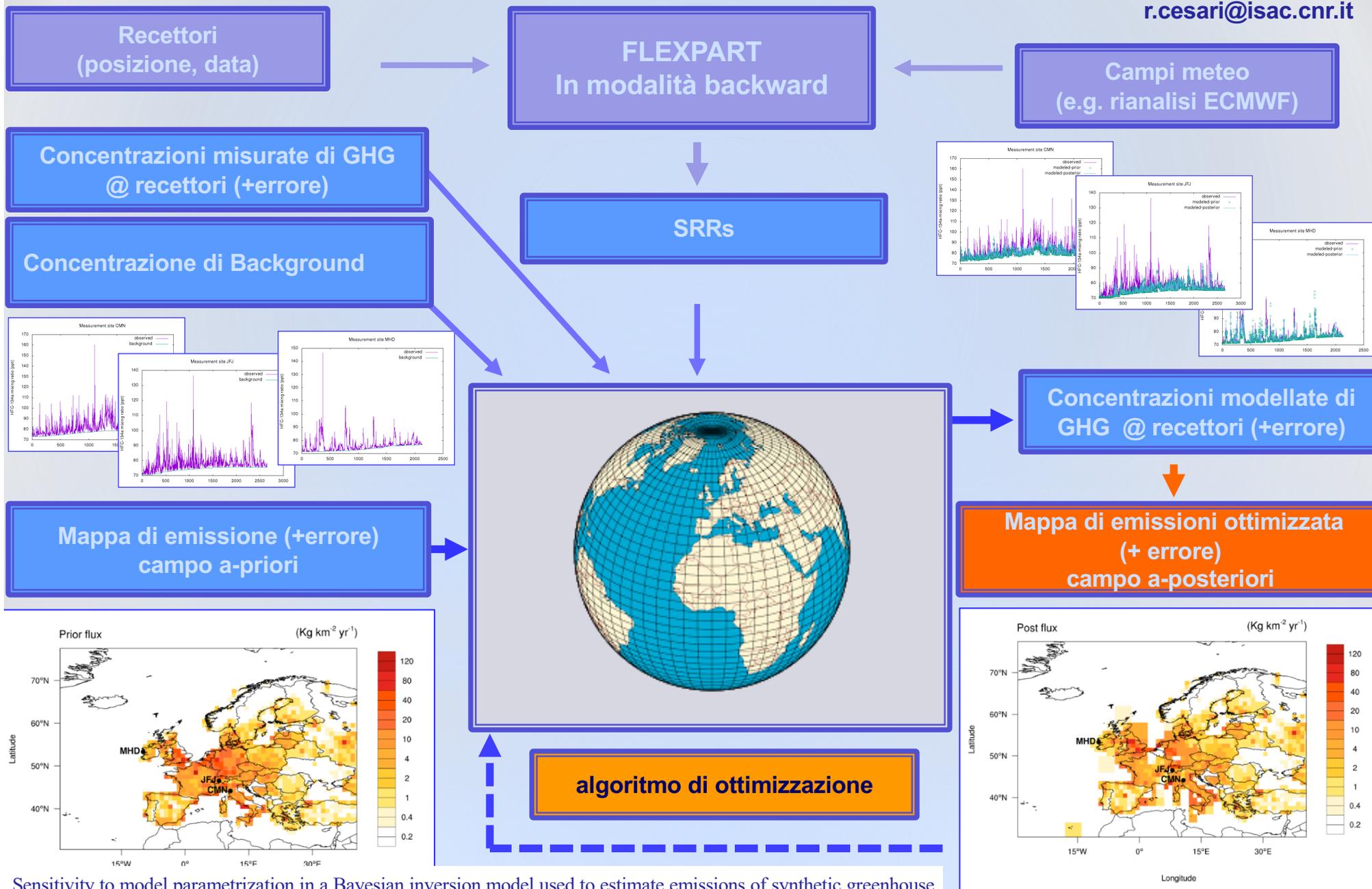
dove \mathbf{x} e \mathbf{x}_0 sono rispettivamente il vettore di stato ottimizzato ed il vettore di stato a priori, \mathbf{B} è la matrice di covarianza dell'errore del campo a priori, \mathbf{H} è l'operatore di trasporto, \mathbf{y} è la concentrazione misurata a cui è sottratto il valore della concentrazione di background, ed \mathbf{R} è la matrice di covarianza dell'errore per le osservazioni.

dove $\mathbf{y}_{(M \times 1)}^{mod}$ è il vettore delle concentrazioni modellate in M punti nello spazio e nel tempo, $\mathbf{x}_{(N \times 1)}$ è il vettore delle N variabili di stato discretizzate nello spazio e nel tempo, $\mathbf{H}_{(M \times N)}$ è l'operatore di trasporto.

Basandosi sul teorema di Bayes, la soluzione più probabile per \mathbf{x} è quella che minimizza la differenza tra le concentrazioni osservate e quelle modellate, tenendo conto delle variabili di stato a priori e delle loro incertezze.

- Receptor oriented models basati sull'analisi delle traiettorie delle masse d'aria

r.cesari@isac.cnr.it

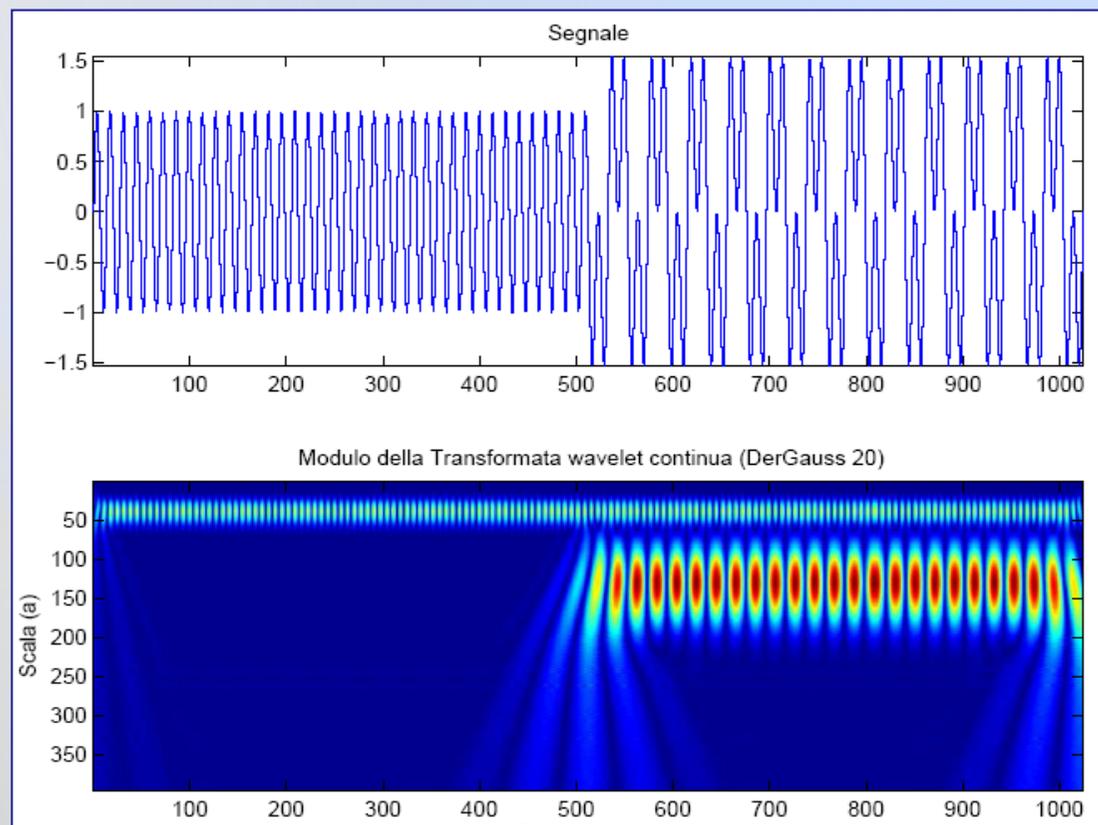


Sensitivity to model parametrization in a Bayesian inversion model used to estimate emissions of synthetic greenhouse gases at the European scale. S. Falasca, R. Cesari, M. Maione, U. Giostra and J. Arduini 2022, *in preparation*

Trasformate Wavelet

Le trasformate Wavelet sono trasformate localizzate sia in tempo che in frequenza, che permettono di analizzare un segnale **localmente** con un dettaglio che dipende dalla scala considerata:

- risoluzione temporale molto buona alle piccole scale
- risoluzione in scala molto buona alle grandi scale.



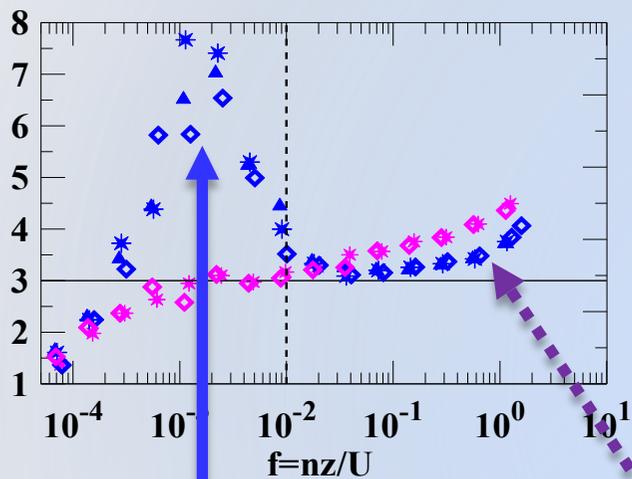
La Trasformata Wavelet individua la distribuzione temporale delle frequenze del segnale

Per la loro caratteristica di **localizzazione**, rappresentano lo strumento più idoneo per studiare segnali **non stazionari** o segnali che presentano caratteristiche differenti alle varie scale o che presentano singolarità (ad esempio, **strutture coerenti** o **fenomeni intermittenti**).

- Trasformate di Wavelet applicate allo studio della turbolenza nel PBL

Applicazioni nello studio del flusso nello strato limite atmosferico: intermittenza della turbolenza a grande e piccola scala

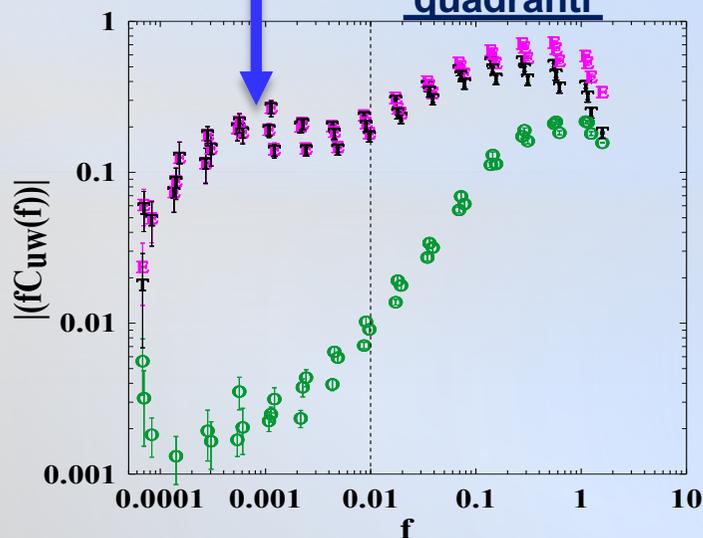
Kurtosis Wavelet



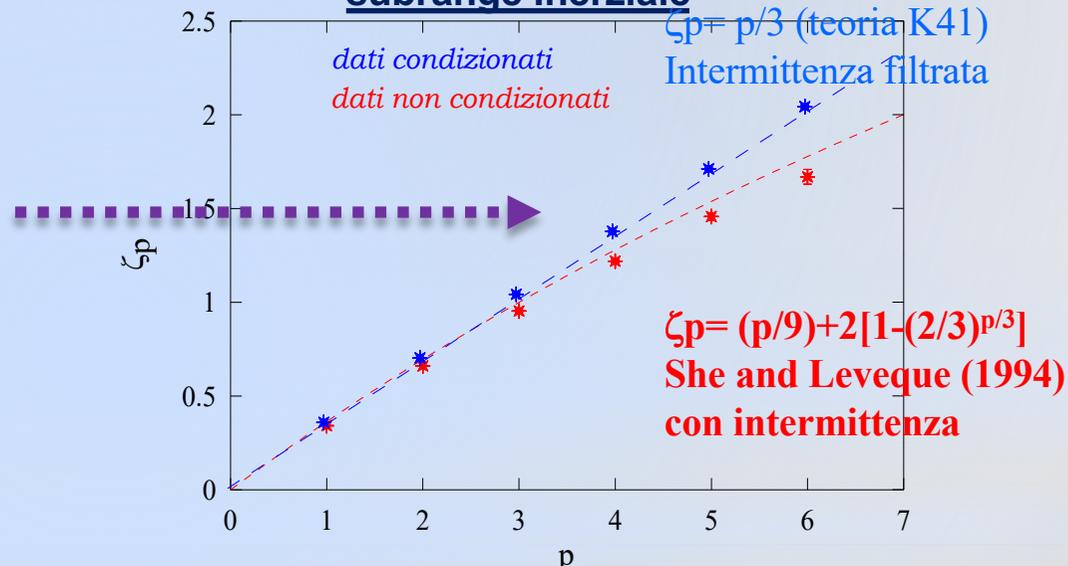
Intermittenza a GRANDE SCALA

Intermittenza a piccola scala

Cospetto Wavelet in 'quadranti'

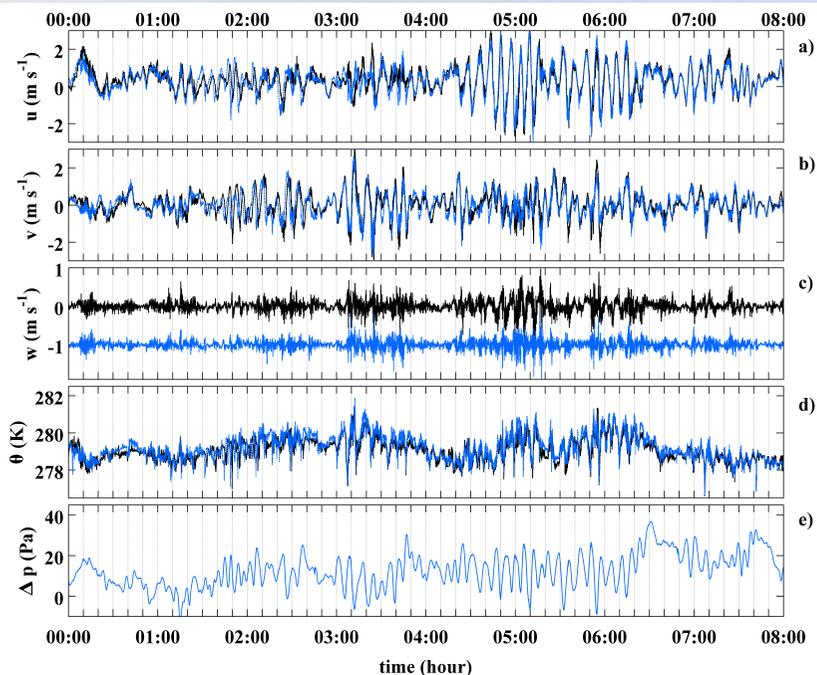


Esponenti di 'scaling' delle funzioni di struttura nel subrange Inerziale

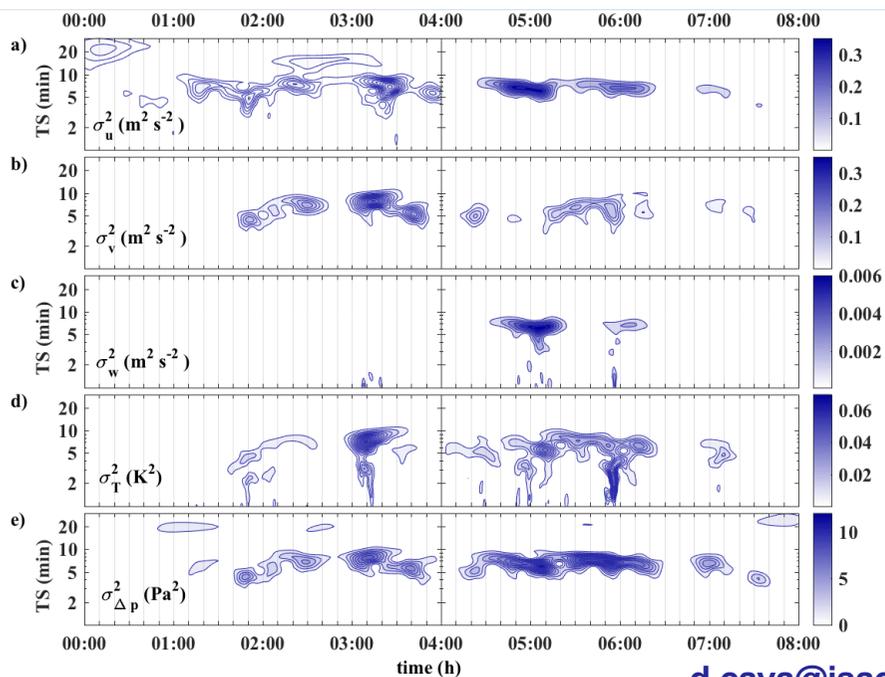
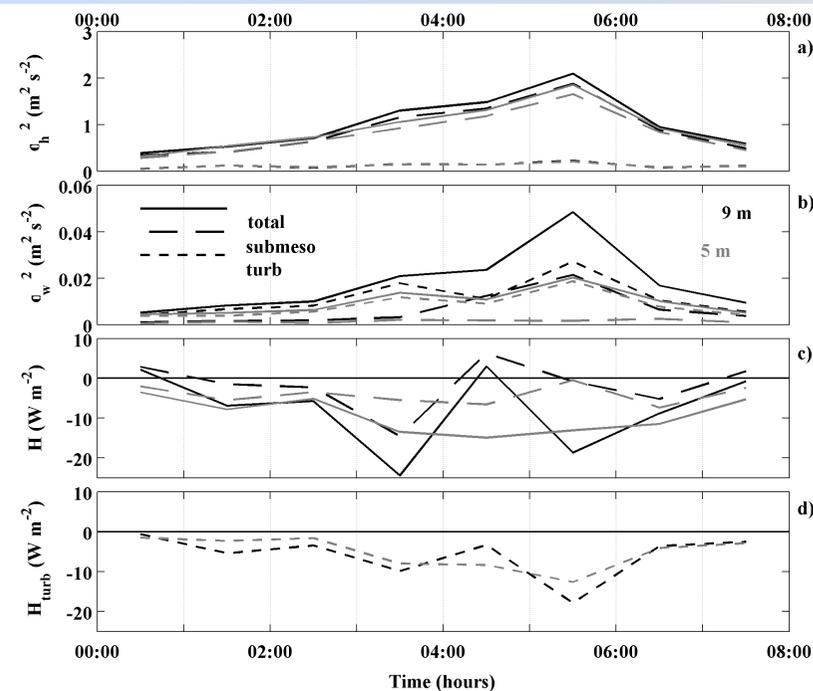


- Cava D, Schipa S, Giostra U. Investigation of low-frequency perturbations induced by a steep obstacle. *Boundary Layer Meteorology*, 115: 27-45, 2005. <https://doi.org/10.1007/s10546-004-2123-y>
- Giostra U, Cava D, Schipa S. Structure functions in a wall-turbulent shear flow. *Boundary Layer Meteorology*, 103: 337-359, 2002. <https://doi.org/10.1023/A:1014917120110>

Trasformate di Wavelet applicate allo studio della turbolenza nel PBL



Applicazioni nello studio del flusso nello strato limite atmosferico: Interazione tra turbolenza e moti a scala più grande in condizioni di forte stabilità atmosferica



Individuazione dell'energia alle varie scale del flusso e scomposizione del segnale per la corretta valutazione del trasporto di energia e massa nello strato limite atmosferico



- **Metodi non asintotici per lo studio delle precipitazioni estreme**

Simplified Metastatistical Extreme Value (SMEV):

f.marra@isac.cnr.it

modello statistico non-asintotico per valori estremi (Marra & al 2019; Marra & al 2020)

Distribuzione degli eventi
ordinari (parent distribution):
 $F(x)$



Distribuzione dei massimi estratti da
campioni di numerosita` n :
 $F(x)^n$

Anziche` assumere $n \rightarrow \infty$ come fatto dal teorema dei valori estremi (Fischer & Tippett 1928), SMEV assume che la 'coda' di $F(x)$ sia nota (ipotesi che puo` essere testata) e descrive la distribuzione dei massimi come:

$$\zeta(x) = F(x)^{\langle n \rangle}$$

Dove $\langle n \rangle$ e` il valore medio della numerosita` degli eventi annuali

SMEV permette di:

- **trattare eventi che emergono da processi diversi** in modo naturale:

$$\zeta(x) = F_1(x)^{\langle n_1 \rangle} \cdot F_2(x)^{\langle n_2 \rangle} \cdot \dots$$

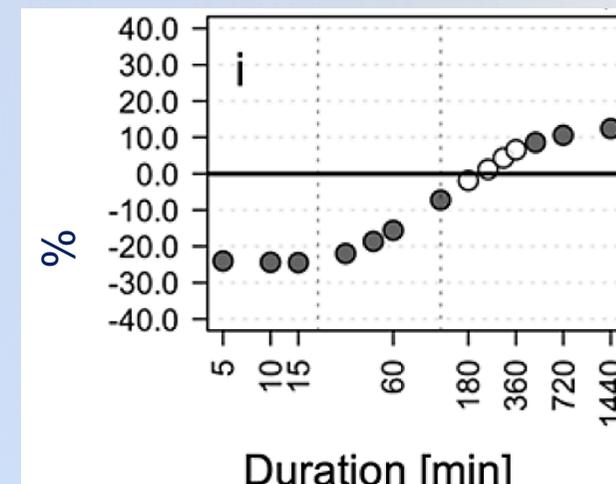
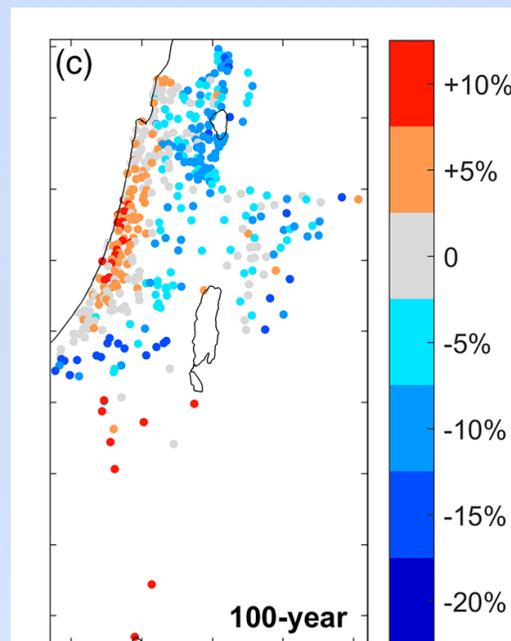
- utilizzare una proporzione di dati molto maggiore (~10-20 volte) rispetto ad approcci basati su estremi, ed avere quindi una **incertezza di stima drasticamente ridotta**

- **Metodi non asintotici per lo studio delle precipitazioni estreme**

Esempi di applicazioni già condotte:

- Mappe di **quantili a elevate tempi di ritorno da dati di telerilevamento** (Rinat et al., 2021; Marra et al., 2022)
- Esplorazione **domande scientifiche altrimenti mascherate da incertezza di stima** (e.g. impatto orografico su precipitazioni di durata breve, statistiche estremi di erosione del suolo e siccità) (Marra et al., 2021a; Formetta et al., 2022, Shmilovitz et al., 2021; Araujo et al., 2022)
- **Analisi e attribuzione degli effetti passati e futuri del cambiamento climatico** (Marra et al., 2021b; Dallan et al., 2022; Amponsah et al., 2022)

Cambiamento previsto per le precipitazioni giornaliere a 100 anni di tempo di ritorno. La stima è basata sui cambiamenti previsti in intensità e accadimento di cicloni mediterranei e altri tipi di eventi (Mediterraneo orientale)



Impatto orografico sulle precipitazioni con 100 anni di tempo di ritorno a diverse durate (provincia di TN)
Formetta et al., 2022

Marra et al., 2021b



References

• Receptor oriented models

Cesari, R., Genga, A., Buccolieri, R., ... Maurizi, A., Ielpo, P. Applied Sciences , 2021, 11(5), pp. 1–26, 2181
E. Andriani, P. Ielpo et al. Current Analytical Chemistry, 2010, 6, 19-25
M. Amodio...P. Ielpo et al. Aerosol and Air Quality Research, 2013: 497–508, 2013

p.ielpo@isac.cnr.it

Cesari D., Merico E. Grasso, F M...Contini D. Atmosphere, 2019, 10(12), 758
Cesari D. ...Contini D. Science of Total Environment 2018, 612, pp. 202–213
Cesari D., Donateo A, Conte M, Contini D. Atmospheric Research 2016, 182, pp. 282–293

d.cesari@isac.cnr.it

Paglione et al., ACP, 2020
Gilardoni et al., ACP, 2014;
Gilardoni et al., PNAS, 2016;

m.paglione@isac.cnr.it

• Source-oriented dispersion models

Cesari, R., Landi, T.C., Maurizi, Int. J. Environment and Pollution, 2019, Vol. 65, Nos. 1/2/3
Cesari, R., Landi, T.C., D'Isidoro, M., Mircea M., Russo F., Malguzzi P., Tampieri F. and Maurizi A., Atmosphere, 2021, 12(2), 192

r.cesari@isac.cnr.it

• Air Trajectories Analysis models

Cesari, R., Paradisi P. and Allegrini P., Int. J. Environ. and Pollution, 2014, Vol. 55, Nos. 1/2/3/4

• Trasformate di Wavelet applicate allo studio della turbolenza nel PBL

Cava D, Mortarini L, Giostra U et al. Quarterly Journal of Royal Meteorological Society, 143: 661-669, 2017.
Cava D, Dias-Júnior CQ, Acevedo O et al. Agricultural and Forest Meteorology, 307 (2022) 10898
Correa PB, Dias-Junior CQ, Cava D et al. . Agricultural and Forest Meteorology, 307, 2021.

d.cava@isac.cnr.it

• Metodi non asintotici per lo studio delle precipitazioni

Araujo DSA, F Marra, C Merow, El Nikolopoulos, 2022.. Env. Res. Lett., 17, 044034,
Dallan E, M Borga, M Zaramella, F Marra, 2022. Geophys. Res. Lett., 49, e2021GL096727.
Marra F, D Zoccatelli, M Armon, E Morin, 2019. Adv. Water Resour., 127, 280-290

f.marra@isac.cnr.it