

Campagne di misura della concentrazione di radon indoor: l'effetto sulle stime aggregate di area di un disegno preferenziale nella scelta dei punti di misura

V° Convegno Nazionale Il Controllo degli Agenti Fisici:
Ambiente, Salute e Qualità della Vita
Novara

Riccardo Borgoni Daniela de Bartolo

7 giugno 2012

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Sommario

- Il campionamento preferenziale
- Stime di area e loro correzioni
- Analisi di simulazione
- Il caso del radon indoor in Lombardia

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Introduzione

Il campionamento di fenomeni indoor presenta un naturale clustering sul territorio legato al concentrarsi degli edifici

Il campionamento ambientale è spesso condotto con disegni preferenziali: si campionano più intensamente aree con concentrazioni attese più elevate per motivi quali

- opportunità
- effetto proporzionale delle concentrazioni
- ...

Disegni campionari così fatti possono indurre distorsioni sulle stime aggregate per aree dovuti alla mancanza di rappresentatività

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Introduzione

Il campionamento di fenomeni indoor presenta un naturale clustering sul territorio legato al concentrarsi degli edifici

Il campionamento ambientale è spesso condotto con disegni preferenziali: si campionano più intensamente aree con concentrazioni attese più elevate per motivi quali

- opportunità
- effetto proporzionale delle concentrazioni
- ...

Disegni campionari così fatti possono indurre distorsioni sulle stime aggregate per aree dovuti alla mancanza di rappresentatività

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Introduzione

Il campionamento di fenomeni indoor presenta un naturale clustering sul territorio legato al concentrarsi degli edifici

Il campionamento ambientale è spesso condotto con disegni preferenziali: si campionano più intensamente aree con concentrazioni attese più elevate per motivi quali

- opportunità
- effetto proporzionale delle concentrazioni
- ...

Disegni campionari così fatti possono indurre distorsioni sulle stime aggregate per aree dovuti alla mancanza di rappresentatività

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

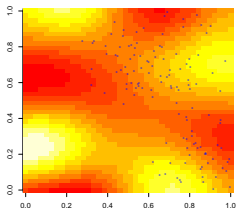
Aggregati di area e loro stime

Studio di simulazione

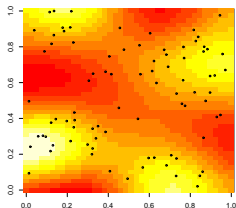
Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Preferenzialità vs Clusterizzazione



Disegno clusterizzato



Disegno preferenziale

- dati simulati con media 10
- media campionaria campionamento clusterizzato: 9.82
- media campionaria campionamento preferenziale: 12.82

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

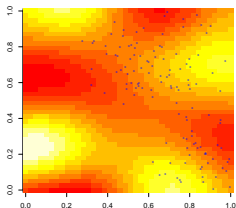
Aggregati di area e loro stime

Studio di simulazione

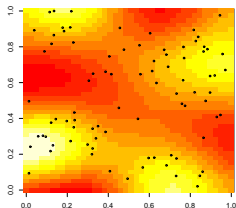
Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Preferenzialità vs Clusterizzazione



Disegno clusterizzato



Disegno preferenziale

- dati simulati con media 10
- media campionaria campionamento clusterizzato: 9.82
- media campionaria campionamento preferenziale: 12.82

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

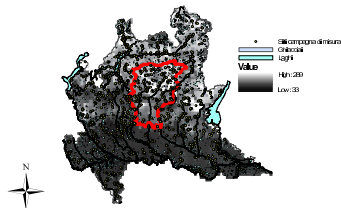
Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

La campagna di ARPA Lombardia

- due campagne negli anni 2003-04 e 2009-2010
- oltre 4600 locali monitorati (3646 nella prima campagna e 966 nella seconda)
- sono stati considerati 2030 locali al pian terreno in edifici destinati ad uso abitativo

Struttura deliberatamente preferenziale nelle due campagne per approfondire aree con una maggiore concentrazione attesa



Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbf{1}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata

$T(D) = T(Y(x_1), \dots, Y(x_n))$ in condizioni di preferenzialità
del campione?

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

**Aggregati di
area e loro
stime**

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbf{1}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata

$T(D) = T(Y(x_1), \dots, Y(x_n))$ in condizioni di preferenzialità
del campione?

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

**Aggregati di
area e loro
stime**

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbf{1}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata

$T(D) = T(Y(x_1), \dots, Y(x_n))$ in condizioni di preferenzialità
del campione?

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

**Aggregati di
area e loro
stime**

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbf{1}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata
 $T(D) = T(Y(x_1), \dots, Y(x_n))$ *in condizioni di preferenzialità del campione?*

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

**Aggregati di
area e loro
stime**

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbb{I}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata
 $T(D) = T(Y(x_1), \dots, Y(x_n))$ *in condizioni di preferenzialità del campione?*

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

**Aggregati di
area e loro
stime**

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbf{1}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata
 $T(D) = T(Y(x_1), \dots, Y(x_n))$ *in condizioni di preferenzialità del campione?*

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

**Aggregati di
area e loro
stime**

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Parametri di area

Il calcolo di statistiche aggregate per aree è necessaria per

- fini descrittivi
- analisi ecologiche
- fini previsionali

Hp: $Y(x)$ è una variabile (aleatoria) regionalizzata su D

Si vogliono prevedere sue caratteristiche aggregate quali:

- $Y(D) = |D|^{-1} \int_D Y(x) dx$
- $P(Y(x) > \tau) = |D|^{-1} \int_D I(x) dx$
dove $I(x) = \mathbf{1}(Y(x) > \tau)$ con τ soglia prefissata

Problema: qual è il significato di una stima aggregata

$T(D) = T(Y(x_1), \dots, Y(x_n))$ in condizioni di preferenzialità
del campione?

Notazione: x_1, \dots, x_n siti campionari, $Y(x_1), \dots, Y(x_n)$ rilevazioni

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Metodo 1: approccio geostatistico - il block kriging

- Stima di $Y(D)$ tramite Block Kriging (BK):

$$\hat{Y}_{BK}(D) = \sum_{i=1}^n \alpha_i Y(x_i) \quad \text{con} \quad \sum_{i=1}^n \alpha_i = 1$$

- previsore BLUP:

$$\operatorname{argmin}_{\alpha_1, \dots, \alpha_n} E(\hat{Y}_{BK}(D) - Y(D))^2 \quad \text{s.v.} \quad \sum_{i=1}^n \alpha_i = 1$$

- si ottiene risolvendo il sistema:

$$\begin{aligned} \sum_{j=1}^n \alpha_j \gamma(x_i - x_j) + \lambda &= \gamma(D, x_i) & i = 1, \dots, n \\ \sum_{i=1}^n \alpha_i &= 1 \end{aligned}$$

$\gamma(h)$: semivariogramma del processo ($Y(x), x \in D$)

$$\gamma(D, x_i) = \frac{1}{|D|} \int_D \gamma(x - x_i) dx$$

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Metodo 2: ponderazione polinomiale e i poligoni di Voronoi

- stima di $Y(D)$ tramite

$$\hat{Y}_w(D) = |D|^{-1} \sum_{i=1}^n w(x_i) Y(x_i) \quad w(x_i) = \pi(x_i)^{-1} \quad i = 1, \dots, n$$

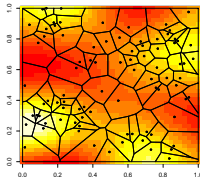
con $\pi(x)$ densità d'inclusione del sito x nel campione

- $\pi(x)$ spesso non è definita a priori e occorre stimarla
- stima utilizzata: l'area del poligono di Voronoi generato dall' i -esima unità campionaria

Poligoni di Voronoi generati dai punti x_1, \dots, x_n

Partizione di D data dai poligoni convessi
 V_1, \dots, V_n dove

$$V_i = \{x : \|x - x_i\| < \|x - x_j\| \quad \forall x_j \neq x_i\} \quad i = 1, \dots, n$$



Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Metodo 3: ricampionamento - l'algoritmo di Olea

- 1 suddividi il campione S in C (cluster set) e U (uncluster set) dove
 $C = \{x : \|x_i - \partial x_i\| < d_0\}$ e $U = S - C$
- 2 sposta $x \in C$ tale che $d(x, \partial x)$ è massima con $\partial x \in U$ in U
- 3 verifica $H_0 : F_U(y) = F_C(y)$ dove $F_a(y)$ è la FR di Y in a con $a = C, U$

Itera 1-3 fino a che H_0 è respinta
Calcola la statistica di sintesi su U

notazione:

- d_0 : mediana delle distanze campionarie
- il test usato per H_0 è Kolmogorov-Smirnov
- ∂x punto più vicino a x

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Studio di simulazione: il disegno sperimentale

Schema della simulazione

1. simula una traiettoria di un processo gaussiano $(Y(x), x \in D)$ e D quadrato di lato unitario con $E(Y(D)) = 10$ e $2\gamma(h) = 30(1 - 0.1/h \sin(h/0.1)), h > 0$
2. simula un campione preferenziale tramite un processo di punto di Cox con intensità $\Lambda(x) = b_0 + b_1 \log Y(x)$
3. calcola gli stimatori di area: media aritmetica, block kriging, correzione poligonale e via ricampionamento
4. iterazione 1-3 per 1000 volte per diverse scelte di b_1
5. calcolo degli indicatori Monte Carlo:
 - $100 \times \sum_{b=1}^{1000} (\bar{T} - 10)/10$ dove $\bar{T} = \sum_{b=1}^{1000} T_b/1000$
 - $V = \sum_{b=1}^{1000} (T_b - \bar{T})^2/1000$
 - $MSE = \sum_{b=1}^{1000} (\bar{T} - 10)^2/1000 = V + (\bar{T} - 10)^2$

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

**Studio di
simulazione**

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Risultati dello studio di simulazione

Campionamento preferenziale

	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$		\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$
$b_1=0.01$				$b_1=0.25$			
V	0.4	0.3	0.4	V	0.6	0.3	0.6
MSE	0.4	0.3	0.4	MSE	1.9	0.4	1.2
RBias (%)	0.1	-0.3	0	RBias (%)	11.5	2.3	7.8
$b_1=0.5$				$b_1=0.75$			
V	1.1	0.6	0.9	V	1.4	1	1.2
MSE	5.5	1.2	3.2	MSE	9.2	3.2	6
RBias (%)	20.9	7.6	15.2	RBias (%)	27.9	14.6	21.9

Algoritmo di Olea (BK con variogramma stimato su U)

	\bar{Y}	$\hat{Y}_{BK}(D)$		\bar{Y}	$\hat{Y}_{BK}(D)$
$b_1=0.5$			$b_1=0.75$		
V	2.0	2.5	V	2.0	2.6
MSE	5.0	5.4	MSE	5.0	5.5
RBias (%)	17.3	17.0	RBias (%)	17.4	17.2

Campionamento clusterizzato

	Distribuzione Uniforme			$b_1 = 0.25$			$b_1 = 0.75$		
	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$
V	0.39	0.34	0.41	0.47	0.32	0.43	1.22	0.58	0.77
MSE	0.39	0.34	0.41	0.47	0.32	0.43	1.22	0.58	0.77
RBias (%)	-0.06	-0.07	-0.08	0.16	0.13	-0.18	-0.07	0.01	-0.07

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Risultati dello studio di simulazione

Campionamento preferenziale

	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$		\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$
$b_1=0.01$				$b_1=0.25$			
V	0.4	0.3	0.4	V	0.6	0.3	0.6
MSE	0.4	0.3	0.4	MSE	1.9	0.4	1.2
RBias (%)	0.1	-0.3	0	RBias (%)	11.5	2.3	7.8
$b_1=0.5$				$b_1=0.75$			
V	1.1	0.6	0.9	V	1.4	1	1.2
MSE	5.5	1.2	3.2	MSE	9.2	3.2	6
RBias (%)	20.9	7.6	15.2	RBias (%)	27.9	14.6	21.9

Algoritmo di Olea (BK con variogramma stimato su U)

	\bar{Y}	$\hat{Y}_{BK}(D)$		\bar{Y}	$\hat{Y}_{BK}(D)$
$b_1=0.5$			$b_1=0.75$		
V	2.0	2.5	V	2.0	2.6
MSE	5.0	5.4	MSE	5.0	5.5
RBias (%)	17.3	17.0	RBias (%)	17.4	17.2

Campionamento clusterizzato

	Distribuzione Uniforme			$b_1 = 0.25$			$b_1 = 0.75$		
	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$
V	0.39	0.34	0.41	0.47	0.32	0.43	1.22	0.58	0.77
MSE	0.39	0.34	0.41	0.47	0.32	0.43	1.22	0.58	0.77
RBias (%)	-0.06	-0.07	-0.08	0.16	0.13	-0.18	-0.07	0.01	-0.07

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Risultati dello studio di simulazione

Campionamento preferenziale

	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$		\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$
$b_1=0.01$				$b_1=0.25$			
V	0.4	0.3	0.4	V	0.6	0.3	0.6
MSE	0.4	0.3	0.4	MSE	1.9	0.4	1.2
RBias (%)	0.1	-0.3	0	RBias (%)	11.5	2.3	7.8
$b_1=0.5$				$b_1=0.75$			
V	1.1	0.6	0.9	V	1.4	1	1.2
MSE	5.5	1.2	3.2	MSE	9.2	3.2	6
RBias (%)	20.9	7.6	15.2	RBias (%)	27.9	14.6	21.9

Algoritmo di Olea (BK con variogramma stimato su U)

	\bar{Y}	$\hat{Y}_{BK}(D)$		\bar{Y}	$\hat{Y}_{BK}(D)$
$b_1=0.5$			$b_1=0.75$		
V	2.0	2.5	V	2.0	2.6
MSE	5.0	5.4	MSE	5.0	5.5
RBias (%)	17.3	17.0	RBias (%)	17.4	17.2

Campionamento clusterizzato

	Distribuzione Uniforme			$b_1 = 0.25$			$b_1 = 0.75$		
	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$	\bar{Y}	$\hat{Y}_w(D)$	$\hat{Y}_{BK}(D)$
V	0.39	0.34	0.41	0.47	0.32	0.43	1.22	0.58	0.77
MSE	0.39	0.34	0.41	0.47	0.32	0.43	1.22	0.58	0.77
Rbias (%)	-0.06	-0.07	-0.08	0.16	0.13	-0.18	-0.07	0.01	-0.07

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Alcuni risultati

La correzione di Voronoi è stata applicata, per alcune province lombarde particolarmente popolate o con concentrazioni attese particolarmente alte, alla stima della:

- concentrazione media di R_n indoor
- probabilità di superamento di soglie predefinite tenendo conto che $\hat{P}(Y(x) > \tau) = n^{-1} \sum_{i=1}^n \mathbf{I}(Y(x_i) > \tau)$

	Bergamo		Brescia	
	corretta	non corretta	corretta	non corretta
Media	174.4	209.4	128.8	169.1
% 400 B_{qm}^{-3}	10.8	14.9	4.0	8.4
% 200 B_{qm}^{-3}	27.2	29.8	15.4	21.9
% 100 B_{qm}^{-3}	49.8	52.9	42.5	47.2

	Sondrio		Milano	
	corretta	non corretta	corretta	non corretta
Media	168.0	183.5	80.8	113.1
% 400 B_{qm}^{-3}	6.0	7.2	1.4	4.2
% 200 B_{qm}^{-3}	19.0	25.9	5.3	12.5
% 100 B_{qm}^{-3}	52.8	61.1	27.0	32.1

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

**Alcuni
risultati sulle
campagne
ARPA
Lombardia**

Conclusioni

Conclusioni

- Il campionamento preferenziale produce distorsione nelle stime di aggregati di area prodotte con le usuali metodologie statistiche e geostatistiche mentre un clustering nelle osservazioni svincolato dai valori della variabile target può indurre solo perdite, più o meno rilevanti, di efficienza della stima;
- In queste circostanze può essere opportuno correggere con adeguati metodi come le ponderazioni basate sui poligoni d'influenza o metodi di ricampionamento (elementi diversi di complessità sono presenti nei due approcci);
- Le metodologie in questione possono essere applicate se le aree d'interesse contengono qualche punto di misura. In caso contrario metodi geostatistici possono risultare più utili sebbene occorra valutare come correggere le procedure per tener conto della preferenzialità.

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Conclusioni

- Il campionamento preferenziale produce distorsione nelle stime di aggregati di area prodotte con le usuali metodologie statistiche e geostatistiche mentre un clustering nelle osservazioni svincolato dai valori della variabile target può indurre solo perdite, più o meno rilevanti, di efficienza della stima;
- In queste circostanze può essere opportuno correggere con adeguati metodi come le ponderazioni basate sui poligoni d'influenza o metodi di ricampionamento (elementi diversi di complessità sono presenti nei due approcci);
- Le metodologie in questione possono essere applicate se le aree d'interesse contengono qualche punto di misura. In caso contrario metodi geostatistici possono risultare più utili sebbene occorra valutare come correggere le procedure per tener conto della preferenzialità.

Sommario

Introduzione

Il problema della preferenzialità

Le campagne di ARPA Lombardia

Aggregati di area e loro stime

Studio di simulazione

Alcuni risultati sulle campagne ARPA Lombardia

Conclusioni

Conclusioni

- Il campionamento preferenziale produce distorsione nelle stime di aggregati di area prodotte con le usuali metodologie statistiche e geostatistiche mentre un clustering nelle osservazioni svincolato dai valori della variabile target può indurre solo perdite, più o meno rilevanti, di efficienza della stima;
- In queste circostanze può essere opportuno correggere con adeguati metodi come le ponderazioni basate sui poligoni d'influenza o metodi di ricampionamento (elementi diversi di complessità sono presenti nei due approcci);
- Le metodologie in questione possono essere applicate se le aree d'interesse contengono qualche punto di misura. In caso contrario metodi geostatistici possono risultare più utili sebbene occorra valutare come correggere le procedure per tener conto della preferenzialità.

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni

Grazie
per l'attenzione

Sommario

Introduzione

Il problema
della
preferenzialità

Le campagne
di ARPA
Lombardia

Aggregati di
area e loro
stime

Studio di
simulazione

Alcuni
risultati sulle
campagne
ARPA
Lombardia

Conclusioni