



## Progetto PULVIRUS

OBIETTIVO 1 - Analisi degli effetti delle misure di distanziamento fisico durante il periodo della pandemia da COVID 19.

ATTIVITÀ 1.4.2 – Analisi dei dati mediante algoritmi di Machine Learning.

Data: 17/10/2022



## ***GRUPPO DI LAVORO***

### **ENEA**

Ilaria D'Elia, Massimo D'Isidoro

### **ISPRA**

Maria Antonietta Reatini

### **Arpa Lombardia**

Andrea Algieri

### **Arpae Emilia Romagna**

Fabiana Scotto



## 1 Sommario

1	Introduzione .....	4
2	Metodologia e dati .....	6
2.1	Dati inquinanti.....	6
2.2	Dati meteorologici.....	7
2.3	Metodologia.....	9
2.3.1	Scelta delle variabili e modello Business-as-usual .....	12
2.3.2	Incertezza delle previsioni .....	14
3	Analisi dei risultati .....	17
3.1	Validazione e valutazione delle prestazioni dei modelli .....	17
3.2	Valutazione dell'incertezza.....	33
3.3	Differenze dell'anno 2020 rispetto all'atteso.....	39
3.3.1	Benzene.....	40
3.3.2	Monossido di Carbonio.....	48
3.3.3	Biossido di azoto.....	55
3.3.4	Ozono .....	62
3.3.5	Particolato atmosferico: PM <sub>10</sub> .....	69
3.3.6	Particolato atmosferico: PM <sub>2,5</sub> .....	76
4	Conclusioni .....	84
5	Bibliografia .....	87
	APPENDICE A Ozono: massimo media mobile di 8 h. ....	89



## 1 INTRODUZIONE

Il progetto Pulvirus nasce per approfondire il legame tra inquinamento atmosferico e diffusione della pandemia, le interazioni fisico-chimiche-biologiche fra polveri atmosferiche e virus, per valutare gli effetti del lockdown sulle concentrazioni atmosferiche degli inquinanti e dei gas serra.

Il 2020 ha offerto un'opportunità senza precedenti di valutare l'effetto della variazione delle emissioni sulla qualità dell'aria, realizzata nel progetto Pulvirus sia attraverso l'analisi delle stazioni di monitoraggio che di sistemi modellistici a scala nazionale. Le misure di lockdown intraprese per fronteggiare la diffusione del virus SARS-CoV-2 hanno determinato una temporanea riduzione delle emissioni degli inquinanti atmosferici e dei gas serra. Nel primo più stringente lockdown, tra la fine di febbraio e di maggio 2020, il settore che ha maggiormente trainato la riduzione è stato il trasporto stradale, con una riduzione delle emissioni degli NO<sub>x</sub> totali di circa il 60%, del PM<sub>2.5</sub> di circa il 66% e di circa l'87% per il CO totale. Il settore industriale ha maggiormente inciso sulla riduzione delle emissioni di SO<sub>x</sub> (circa 90%) e COVNM (circa 80%). Il settore residenziale/terziario ha registrato un incremento delle emissioni di PM<sub>2.5</sub>, legato alla maggiore presenza delle persone nelle abitazioni. Il settore marittimo ha poi contribuito ad una riduzione delle emissioni di NO<sub>x</sub> di circa l'8% e di SO<sub>x</sub> di circa il 3%. Per maggiori dettagli si rimanda ai report dell'Obiettivo 2 del progetto Pulvirus, disponibili al link seguente <https://www.pulvirus.it/index.php/documentazione-obiettivo-2/>. La valutazione degli effetti delle misure intraprese durante il lockdown, e per l'intero 2020, attraverso l'analisi delle concentrazioni degli inquinanti rilevate dalle stazioni di monitoraggio italiane risulta particolarmente difficile per i complessi meccanismi che regolano diffusione, trasporto, trasformazione e deposizione di sostanze inquinanti. In tale contesto, la principale finalità dell'Obiettivo 1 del Progetto Pulvirus è la quantificazione, al netto della variabilità indotta dalle condizioni meteorologiche, della variazione della concentrazione dei principali inquinanti determinata dai provvedimenti adottati per ridurre la diffusione del contagio, utilizzando come base i dati rilevati dalle stazioni di monitoraggio nel corso del 2020, e, per confronto, le serie storiche degli anni precedenti.



In particolare, l'attività 1.4.2 ha avuto come finalità l'analisi dei dati tramite l'applicazione di modelli di Machine learning (ML), ed in particolare della tecnica "Random Forest" (RF). Per quantificare le riduzioni, indotte dalle misure di lockdown, abbiamo confrontato i valori stimati business-as-usual che sarebbero stati osservati nel 2020 in assenza del lockdown, nelle condizioni meteorologiche reali, in ogni stazione di misurazione, con i livelli di concentrazioni effettivamente osservati.

Per lo sviluppo è stata utilizzata la libreria "Ranger" (Wright MN 2017) di R (R Core Team 2021).



## 2 METODOLOGIA E DATI

### 2.1 Dati inquinanti

I dati di concentrazione dei vari inquinanti in atmosfera, misurati nel corso degli anni nelle stazioni di monitoraggio distribuite sul territorio nazionale, sono raccolti e archiviati in ISPRA nel database InfoARIA. I dati sono trasmessi dalle Regioni e Province Autonome secondo quanto previsto dalla Decisione 2011/850/EU.

Ai fini del progetto sono stati utilizzati i dati disponibili attraverso questa fonte per il periodo 2013 al 2019, soddisfacenti i criteri di completezza della serie storica a livello mensile, annuale e dell'intero periodo selezionato (2013-2019) riportati al link seguente <https://progettopulvirus.github.io/pulviruS/completezza.html>

Poiché i dati relativi al 2020 sono stati trasmessi nel 2021 e il processo di trasmissione si è concluso a ottobre 2021, per poter avviare le analisi già nel corso della seconda metà del 2020, si è proceduto al reperimento dei dati parziali (relativi al primo semestre 2020) seguendo due canali:

- sinergia con le ARPA partecipanti al progetto LIFE Prepair (Piemonte, Valle d'Aosta, Lombardia, Trento, Veneto, Friuli Venezia Giulia ed Emilia-Romagna), per evitare richieste duplici degli stessi dati. I colleghi di ARPAE che curano tale attività nell'ambito del progetto LIFE Prepair hanno messo a disposizione i dati raccolti per il periodo di studio.
- Richiesta ai componenti della rete dei referenti sulla qualità dell'aria dell'SNPA (RR-TEM II/02) per le restanti 14 ARPA/APPA.

Nel mese di ottobre 2021 sono poi stati raccolti e completati i dataset con i dati relativi all'intero 2020 utilizzando il database InfoARIA.

Gli inquinanti considerati sono stati: PM<sub>10</sub>, PM<sub>2.5</sub>, NO<sub>2</sub>, O<sub>3</sub>, C<sub>6</sub>H<sub>6</sub>, CO.

I dati di base sono medie orarie per NO<sub>2</sub>, O<sub>3</sub>, C<sub>6</sub>H<sub>6</sub>, CO.

Per PM<sub>10</sub> e PM<sub>2.5</sub> in larga prevalenza (oltre l'80%) i dati di base sono medie giornaliere. Sono presenti, comunque, un certo numero di serie di dati con risoluzione temporale oraria e bioraria.

Ai fini delle successive analisi i dati sono stati aggregati su base giornaliera (medie giornaliere). Il criterio di aggregazione prescelto prevede la disponibilità di almeno 18 record orari al giorno



(copertura giornaliera del 75%) per considerare valida una media giornaliera (ovvero 9 record nel caso di dati biorari). Nel caso di copertura insufficiente la relativa media giornaliera risulta come *missing* (e riportata come “NA”).

Per quanto riguarda l’ozono sono stati calcolati due parametri diversi: la media oraria massima giornaliera e la massima media mobile di otto ore giornaliera, in quanto indicativi maggiormente dell’esposizione a questo inquinante.

## 2.2 Dati meteorologici

I dati meteorologici sono stati estratti dal set di dati *ERA5* generato dal servizio Copernicus sui cambiamenti climatici. Copernicus è un programma di osservazione della Terra frutto della collaborazione tra Agenzia Spaziale Europea e Commissione Europea. I dati ERA5, attualmente disponibili per uso pubblico, sono dati spaziali di rianalisi di condizioni meteorologiche orarie, ovvero dati derivanti da combinazioni di modelli meteorologici con dati osservativi di satelliti e di sensori terrestri, con l’obiettivo di costruire una banca dati meteorologica coerente ed omogenea a partire dal 1979 fino ad oggi, che si aggiorna con cadenza mensile. In particolare, per tutte le variabili selezionate, ad eccezione della stima dell’altezza del planetary boundary layer (pbl) non contenuta in questo dataset, i dati sono stati estratti da *ERA-land (terraferma)*; questi hanno risoluzione temporale oraria (UTC) e sono disponibili dal 1981 ad oggi, su tutto il globo terrestre. Sono dati a risoluzione orizzontale (latitudine x longitudine) pari a  $0.1^\circ \times 0.1^\circ$ . La fonte (chiamata Fonte 1 in tabella 2) è la seguente:

<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview>).

Per la sola pbl sono stati usati i dati *ERA5 single-levels* (disponibili dal 1979 ad oggi, su tutto il globo terrestre, sempre con risoluzione oraria (UTC)). Sono in questo caso dati a risoluzione (latitudine x longitudine) pari a  $0.25^\circ \times 0.25^\circ$ . La fonte (chiamata Fonte 2 in tabella 2) è la seguente:

<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form>

Tab. 1: Variabili meteo di origine scelte (nomi assegnati in questo Progetto); dati orari.

Variabile	Descrizione	Unità di misura	Metodo di misurazione	Fonte
<i>temp</i>	temperatura dell'aria a 2 metri d'altezza sulla superficie terrestre	K (gradi Kelvin)	ad ogni ora viene riportato il valore relativo all'ora precedente	Fonte 1
<i>dewP</i>	temperatura del punto di rugiada (o di dewpoint) dell'aria a 2 metri d'altezza sulla superficie terrestre	K (gradi Kelvin)	ad ogni ora viene riportato il valore relativo all'ora precedente	Fonte 1
<i>prec</i>	precipitazione totale sulla superficie terrestre	m (metri)	ad ogni ora viene riportato il valore cumulato dall'inizio del giorno fino a quell'ora	Fonte 1
<i>winU</i>	componente orizzontale verso Est del vento a 10 metri d'altezza sulla superficie terrestre	m/s (metri al secondo)	ad ogni ora viene riportato il valore relativo all'ora precedente	Fonte 1
<i>winV</i>	componente orizzontale verso Nord del vento a 10 metri d'altezza sulla superficie terrestre	m/s (metri al secondo)	ad ogni ora viene riportato il valore relativo all'ora precedente	Fonte 1
<i>pres</i>	pressione dell'aria sulla superficie terrestre	Pa (Pascal)	ad ogni ora viene riportato il valore relativo all'ora precedente	Fonte 1
<i>Srad</i>	radiazione solare netta che raggiunge un piano orizzontale sulla superficie terrestre	J/m <sup>2</sup> (Joule su metri quadri)	ad ogni ora viene riportato il valore cumulato dall'inizio del giorno fino a quell'ora	Fonte 1
<i>pbl</i>	altezza dello strato limite planetario	m (metri)	ad ogni ora viene riportato il valore relativo all'ora precedente	Fonte 2

Ad ogni stazione di misura di qualità dell'aria è stato associato il punto del grigliato di ERA5 più vicino con valore non mancante, in base alle coordinate spaziali, e sono quindi stati associati i corrispondenti valori delle variabili meteo di origine del grigliato di ERA5. Le variabili meteo di origine sono state ridefinite per l'ottenimento di quelle di interesse, nel modo di seguito descritto (variabili meteo di output). I valori di queste variabili sono stati infine calcolati su base giornaliera, con data e ora nel sistema orario solare italiano per tutto l'anno, per essere successivamente interfacciati con i corrispondenti dati di qualità dell'aria.

Tab. 2: Variabili meteo di output (nomi assegnati in questo Progetto); dati giornalieri.

Variabile	Descrizione	Unità di misura	Metodo di misurazione
<i>t2m</i>	temperatura media dell'aria a 2 metri d'altezza sulla superficie terrestre	°C	Media giornaliera dei valori di <i>temp</i>
<i>tmin2m</i>	temperatura minima dell'aria a 2 metri d'altezza sulla	°C	Minima giornaliera dei valori



Variabile	Descrizione	Unità di misura	Metodo di misurazione
	superficie terrestre		di <i>temp</i>
<i>tmax2m</i>	temperatura massima dell'aria a 2 metri d'altezza sulla superficie terrestre	°C	Massima giornaliera dei valori di <i>temp</i>
<i>rh</i>	umidità relativa percentuale dell'aria sulla superficie terrestre	%	Media giornaliera ottenuta da elaborazione dei valori orari di <i>temp</i> e <i>dewP</i>
<i>tp</i>	precipitazione cumulata sulla superficie terrestre	mm	Cumulato giornaliero al giorno g ottenuto da elaborazione dei valori di <i>prec</i>
<i>ptp</i>	precipitazione cumulata sulla superficie terrestre	mm	Cumulato giornaliero al giorno g-1 ottenuto da elaborazione dei valori di <i>prec</i>
<i>u10m</i>	componente orizzontale verso Est del vento a 10 metri d'altezza sulla superficie terrestre	m/s	Media giornaliera dei valori di <i>winU</i>
<i>v10m</i>	componente orizzontale verso Nord del vento a 10 metri d'altezza sulla superficie terrestre	m/s	Media giornaliera dei valori di <i>winV</i>
<i>wspeed</i>	Intensità della velocità orizzontale del vento a 10 metri d'altezza sulla superficie terrestre	m/s	Media giornaliera al giorno g ottenuta da elaborazione dei valori orari di <i>winU</i> e <i>winV</i>
<i>pwspeed</i>	Intensità della velocità orizzontale del vento a 10 metri d'altezza sulla superficie terrestre	m/s	Media giornaliera al giorno g-1 ottenuta da elaborazione dei valori orari di <i>winU</i> e <i>winV</i>
<i>wdir</i>	direzione di provenienza del vento orizzontale a 10 metri d'altezza sulla superficie terrestre, espressa in settori di prevalenza di direzione (i settori sono espressi con numeri da 1 a 8, in base alla suddivisione angolare del piano in 8 quadranti da 45°, contati in senso orario dal Nord)	-	Settori di maggioranza relativa giornaliera ottenuti da elaborazione dei valori orari di <i>winU</i> e <i>winV</i>
<i>sp</i>	pressione dell'aria sulla superficie terrestre	hPa	Media giornaliera dei valori di <i>pres</i>
<i>nirradiance</i>	irradianza solare netta che raggiunge un piano orizzontale sulla superficie terrestre	W/m <sup>2</sup>	Cumulato giornaliero ottenuto da elaborazione dei valori di <i>Srad</i>
<i>pbl00</i>	altezza dello strato limite planetario alle ore 00	km	Singolo valore orario dei valori di <i>pblH</i>
<i>pbl12</i>	altezza dello strato limite planetario alle ore 12	km	Singolo valore orario dei valori di <i>pblH</i>
<i>pblmin</i>	altezza minima dello strato limite planetario	km	Minima giornaliera dei valori di <i>pblH</i>
<i>pblmax</i>	altezza massima dello strato limite planetario	km	Massima giornaliera dei valori di <i>pblH</i>

## 2.3 Metodologia

Negli ultimi anni l'uso di tecniche di ML è cresciuto rapidamente a causa, in parte, della disponibilità di 'big data' e di una maggiore potenza di calcolo. Il ML supervisionato si occupa dello sviluppo di



algoritmi e tecniche in grado di apprendere informazioni di qualsivoglia natura dai dati e di utilizzare la conoscenza appresa per la sintesi di nuova conoscenza, sia essa sotto forma di una classificazione o della previsione di un risultato. Le tecniche di ML sono non parametriche e hanno il grande vantaggio di non dover fare assunzioni come la normalità del campione, l'omoschedasticità, l'indipendenza, l'aderenza ad altre rigorose ipotesi parametriche e l'attenta gestione degli effetti di interazione che sono invece richieste per i modelli statistici.

La tecnica RF, introdotta da Breiman (Breiman, Random forests 2001), è un algoritmo di apprendimento supervisionato di tipo *ensemble* ampiamente usata nelle analisi di previsione di sistemi ecologici complessi. L'ensemble learning usa differenti algoritmi derivanti dal ML e dalla statistica per poi combinarli tra loro e ottenere un modello finale (*ensemble model*) dalla potenza predittiva maggiore rispetto ai singoli modelli di partenza (*base learners*). La RF si avvale del *bagging* (*Bootstrap AGGREGatING*) (Breiman, Bagging predictors machine learning 1996) come metodo di ensemble learning che si basa sulla media delle previsioni. Il *bagging* è un metodo per generare più versioni di un predittore e utilizzarle per ottenere un predittore aggregato a partire da sottoinsiemi casuali del set di addestramento originale (campione bootstrap). Mentre il *bagging* è una tecnica generale, che in linea di principio può essere utilizzata per ridurre la varianza di qualsiasi modello base, la RF presuppone che questi stimatori di base siano dati da alberi di classificazione o regressione. Un albero di decisione (CART - *Classification And Decision Tree*), metodo definito da Breiman (Breiman, Friedman e Olshen, Classification and Regression Trees 1984), è un modello predittivo, dove ogni nodo interno rappresenta una variabile, un arco verso un nodo figlio rappresenta un possibile valore per quella variabile e una foglia il valore predetto per la variabile di risposta. Un albero di regressione è costruito partizionando ricorsivamente il campione (= il "nodo radice") in gruppi sempre più omogenei fino ai "nodi terminali". Ogni suddivisione si basa sui valori di una variabile ed è selezionata secondo un criterio di suddivisione. Nella tecnica del RF, oltre ad utilizzare un campione *bootstrap* come insieme di stima, ogni albero viene fatto crescere selezionando, ad ogni nodo, un sottoinsieme casuale delle variabili esplicative disponibili; questo aggiunge un'ulteriore casualità durante la costruzione di ciascun albero, che consente di ridurre la correlazione tra i vari alberi stimati e, di conseguenza, la varianza del modello finale.



Nel RF non viene eseguita alcuna fase di “potatura”, quindi tutti gli alberi della foresta sono alberi massimali.

Il predittore basato su RF può essere definito come:

$$\hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{b=1}^B T_b(x; \Theta_b)$$

Dove  $\Theta_b$  caratterizza il b-mo albero della foresta casuale in termini di variabili di split, punto di split in ogni nodo, valori del nodo terminale;  $T_b(x; \Theta_b)$  è la previsione del b-mo albero della foresta casuale. Grazie alla flessibilità e alle elevate capacità predittive, l’algoritmo RF sta avendo una rapida diffusione nelle applicazioni in campo ambientale ( (Rahman, et al. 2020), (Araki, Shima e Yamamoto 2018), (Stafoggia, et al. 2019)). Inoltre, è molto facile da usare, nel senso che ha solo due parametri, il numero di variabili nel sottoinsieme casuale ad ogni nodo (mtry) e il numero di alberi nella foresta (n\_trees).

Nel nostro caso per determinare i valori ottimali dei parametri, sono stati provati più modelli con diverse combinazioni. I parametri sono stati scelti in base ai migliori risultati di performance del modello. Per il numero di variabili nel sottoinsieme casuale si è scelto, ad ogni nodo, di seguire la letteratura, che indica per la regressione mtry = d/3, dove ‘d’ sono le variabili disponibili. Questo risultato deriva da diverse simulazioni progettate da Breiman (Breiman, Random forests 2001) nell'articolo seminale. In particolare, nel nostro studio i parametri utilizzati sono mtry=5 e n\_trees=1000.



### 2.3.1 Scelta delle variabili e modello Business-as-usual

Un primo modello, con tutte le variabili indipendenti disponibili, ha consentito di selezionare, per il modello finale, le *feature* con un indice di importanza maggiore (Genuer, Poggi e Tuleau-Malot 2010).

Le variabili meteorologiche considerate per il presente studio sono intrinsecamente inter-dipendenti. Per diminuire le distorsioni nelle metriche di importanza con insiemi di dati che contengono caratteristiche collineari sono stati addestrati più modelli escludendo ricorsivamente le variabili maggiormente correlate e selezionando poi le *features* che hanno raggiunto i migliori punteggi di performance. Le variabili sono state selezionate in base al contributo in termini di aumento del coefficiente di determinazione ( $R^2$ ). Lo svantaggio di questo approccio è che è abbastanza impegnativo dal punto di vista computazionale, dato che richiede ripetutamente il riaddestramento del modello.

A titolo di esempio per selezionare la migliore variabile tra quelle disponibili, nel caso della temperatura:

- "t2m", temperatura media (2m)
- "tmin2m", temperatura minima (2m)
- "tmax2m", temperatura massima (2m)

sono stati addestrati più modelli RF tenendo fuori 2 variabili per volta e calcolando ogni volta  $R^2$  medio e l'errore quadratico (MSE) medio. Nella Tab. 3 sono riportati i risultati.

Tab. 3 -  $R^2$  medio e MSE medio dei RF che utilizzavano rispettivamente la temperatura media, minima e massima

	"t2m"	"tmin2m"	"tmax2m"
<b><math>R^2</math></b>	0.5777692	0.5734996	0.5753786
<b>MSE</b>	93.01554	93.98849	93.63695

La variabile migliore anche se per pochi decimali è risultata essere la temperatura media.

Di seguito nella

il grafico a barre con le variabili selezionate e utilizzate per tutti gli inquinanti, in quanto è stato verificato tramite il confronto di opportuni indici di performance, che le variabili che davano il contributo maggiore erano sostanzialmente le stesse per tutti gli inquinanti. Nel grafico le variabili sono ordinate tramite la media dell'indice di importanza; la media per ogni variabile è raffigurata da un triangolo.

Le variabili sono state selezionate in base al contributo in termini di aumento di  $R^2$

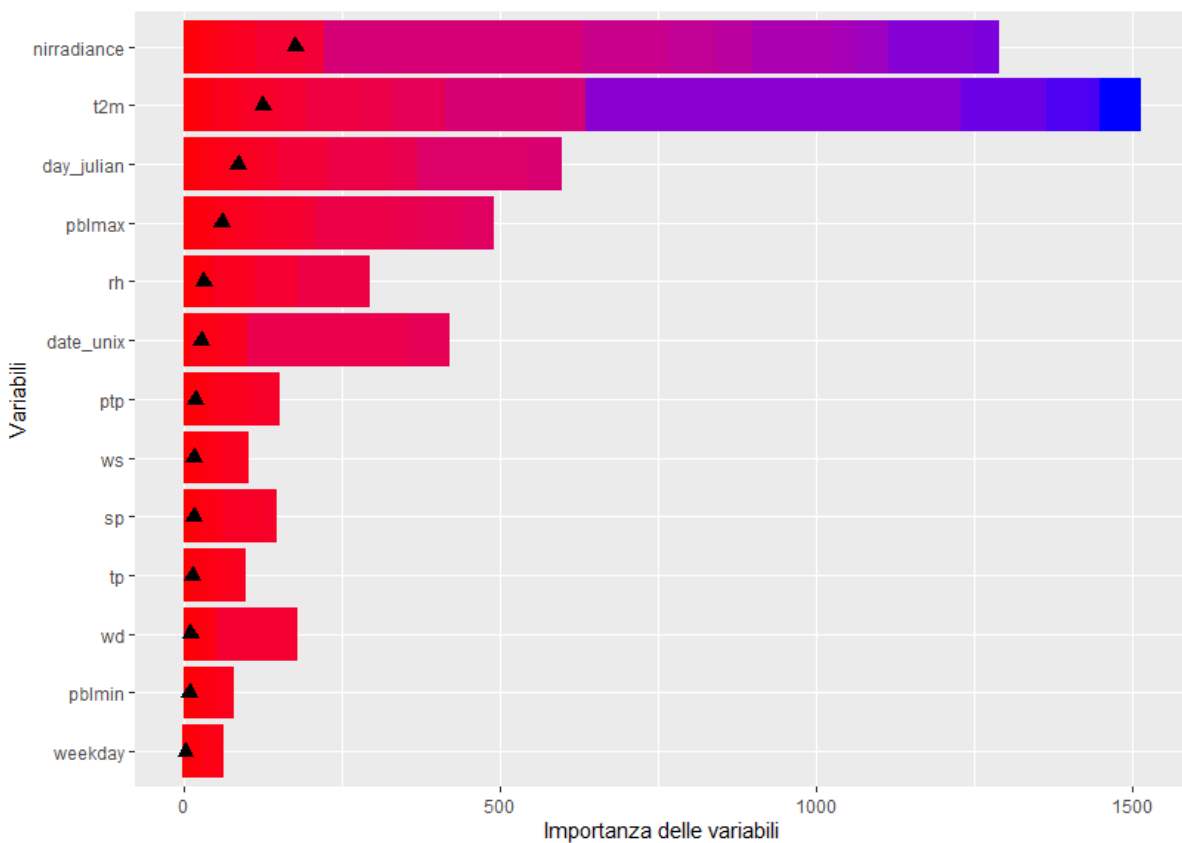


Fig. 1 – Grafico a barre dell'importanza delle variabili per tutti i modelli degli inquinanti addestrati (Benzene, PM<sub>10</sub>, PM<sub>2.5</sub>, Monossido di carbonio, Ozono, Biossido di azoto)

Le variabili meteorologiche selezionate con questa metodologia sono:



- Pressione
- Temperatura media (2m)
- Precipitazione cumulata
- Precipitazione cumulata del giorno precedente
- Umidità relativa
- Irradianza solare netta
- Altezza dello strato limite planetario (minimo giornaliero)
- Altezza dello strato limite planetario (massimo giornaliero)
- Direzione del vento
- Velocità del vento

Inoltre, sono stati utilizzati un termine di tendenza sotto forma di data Unix (numero di secondi dal 1° gennaio 1970), e due termini stagionali sotto forma di giorno giuliano (giorno dell'anno) e giorno della settimana (S. Grange, et al. 2018), (Grange e Carslaw, Using meteorological normalisation to detect interventions in air quality time series 2019)).

Per quantificare le differenze medie di concentrazione al netto della meteorologia riscontrate nel 2020 (a partire dall'inizio del lockdown) rispetto agli anni precedenti, è stato preliminarmente generato un dataset di concentrazioni medie giornaliere rappresentativo di una situazione “*Business As Usual*” ovvero corrispondente a una situazione ipotetica in cui le restrizioni dovute alle misure di contenimento della pandemia non fossero state implementate. A tal fine sono stati addestrati gli algoritmi RF su misurazioni di inquinanti e parametri meteorologici degli anni precedenti (Petetin, et al. 2020), (Grange, Lee, et al. 2021).

### 2.3.2 Incertezza delle previsioni

L'incertezza delle previsioni delle random forests, per problemi di tipo regressivo, può essere stimata usando diversi approcci:



- Il metodo Jackknife-after-Bootstrap (Wager, Hastie e Efron 2014)
- L'approccio U-statistics di (Mentch e Hooker 2016)
- L'approccio delle simulazioni Monte Carlo (sia variabile target che covariate) di (Coulston, et al. 2016)
- Il metodo Quantile Regression Forests (QRF) (Meinshausen e Ridgeway 2006)

Gli approcci di (Wager, Hastie e Efron 2014) e (Mentch e Hooker 2016) stimano gli errori standard dei valori attesi delle previsioni, utilizzati poi per costruire gli intervalli di confidenza, mentre gli approcci di (Coulston, et al. 2016) e (Meinshausen e Ridgeway 2006) stimano gli intervalli di previsione. Questi risultati forniscono un mezzo con cui l'incertezza nelle previsioni delle random forests può essere formalmente quantificata. Queste stime della varianza possono essere utilizzate in un quadro di test standard per valutare le ipotesi sulla significatività predittiva di una variabile. Poiché le foreste casuali sono un metodo non parametrico, non c'è un test della significatività di un particolare parametro, come avviene nei modelli parametrici standard, ad esempio la regressione lineare. Nelle random forests le previsioni stesse prendono il posto di questi parametri convenzionali.

Nel nostro studio abbiamo applicato l'approccio di (Wager, Hastie e Efron 2014) per stimare la variabilità delle previsioni e calcolare l'intervallo di confidenza al 95% come segue:

$$\text{valore previsto} \pm 1.96 * \text{standard error}$$

È importante sottolineare che tali intervalli di confidenza sono validi e centrati sulla previsione attesa da una foresta casuale e non necessariamente sul vero valore della funzione di regressione sottostante. Individuato l'intervallo di confidenza possiamo ritenere significativi, ai fini della nostra indagine, tutti quei valori osservati che “non ricadono” in tale intervallo e che hanno quindi una variazione, in positivo o negativo, maggiore della variabilità attesa della previsione.

L'algoritmo seguito per quantificare la significatività della variazione media mensile delle concentrazioni degli inquinanti riscontrata, ovvero le differenze tra il dato previsto dal modello e le concentrazioni osservate da marzo a dicembre è stato il seguente: per ogni giorno del periodo di lockdown è stato valutato se il valore osservato ricadeva all'interno o fuori dall'intervallo di



confidenza. In base al numero di volte in cui un valore osservato è risultato significativamente differente da quello predetto è stata calcolata la percentuale mensile di significatività. I valori così calcolati sono riportati nel paragrafo di dettaglio dell'inquinante





### 3 ANALISI DEI RISULTATI

#### 3.1 Validazione e valutazione delle prestazioni dei modelli

Il processo per stabilire se i risultati di un modello stimato siano accettabili è noto come validazione. Per valutare se un algoritmo di apprendimento supervisionato ha un buon grado di generalizzazione, ed è quindi in grado di predire nuove istanze, è auspicabile utilizzare un set di dati che non è stato utilizzato nell'addestramento. A tale scopo abbiamo diviso il dataset delle concentrazioni in due insiemi:

- un training set, usato nella fase di addestramento per consentire al modello di apprendere la relazione nascosta tra i dati;
- un testing set, usato nella fase di validazione per valutare le prestazioni del modello.

In condizioni ideali, il modello funzionerà ugualmente bene su entrambe le porzioni di dati.

Nel nostro studio abbiamo utilizzato come set di training i dati compresi nel periodo dal 1° gennaio 2013 al 31 dicembre 2019 e come testing set, per la validazione, il periodo che va dal 1° gennaio 2020 al 14 febbraio 2020. I modelli addestrati e validati sono stati utilizzati per creare le stime "Business As Usual" dal 15 febbraio 2020 fino a dicembre 2020 da confrontare con i dati effettivamente osservati nello stesso periodo. Per valutare l'accuratezza delle previsioni del modello abbiamo calcolato:

- il *Mean Square Error (MSE)*, ovvero la media aritmetica dei quadrati degli scarti tra il valore predetto e l'osservato la cui formula è la seguente:

$$MSE = \frac{\sum_{i=1}^n (O_i - P_i)^2}{n}$$

dove  $P_i$  è la concentrazione prevista dal modello per l'osservazione  $i$ -esima,  $O_i$  la concentrazione osservata (misurata) e  $n$  il numero di osservazioni. Le RF forniscono una stima



dell'errore di test (MSE) per il set di dati Out Of Bag (OOB). Lo stesso indice è stato calcolato sul testing set.

- Il Normalized Mean Square Error (NMSE) definito come il rapporto tra il MSE e il prodotto della media dei valori osservati e la media dei valori predetti; tale indice evidenzia la dispersione nel set di dati. La formula è la seguente:

$$NMSE = \frac{MSE}{\bar{P} * \bar{O}}$$

La normalizzazione per il prodotto dei valori osservati e i valori predetti assicura che l'NMSE non sia distorto verso modelli che prevedono troppo o troppo poco. Valori più piccoli di NMSE denotano una migliore performance del modello.

- $R^2$  definito anche varianza spiegata o coefficiente di determinazione indice restituito in output dal pacchetto R Ranger utilizzato per l'algoritmo di RF; tale indice è calcolato sul campione *out-of-bag sia del set di training (tr\_r\_squared\_oob) che del set di testing (cv\_r\_squared\_oob)*.

I campioni OOB sono quei campioni che non sono inclusi nei campioni bootstrap. I campioni OOB sono essenzialmente set di convalida generati dal processo bootstrap e che non coincidono con il set di validazione scelto nello studio (1° gennaio 2020 al 14 febbraio 2020). Un'osservazione importante è che l'errore OOB viene calcolato anche durante l'addestramento della foresta casuale, pertanto nel nostro studio abbiamo calcolato il NMSE sia sul campione out-of-bag del periodo di training (1° gennaio 2013 - 31 dicembre 2019) e che indicheremo con NMSE\_t\_oob sia il NMSE sul set di validazione scelto che indicheremo con NMSE\_cv.

Nei grafici seguenti, per ciascun inquinante, sono riportati i box plot delle distribuzioni del Normalized Mean Square Error per tipo di stazione e tipo di zona della stazione di monitoraggio.

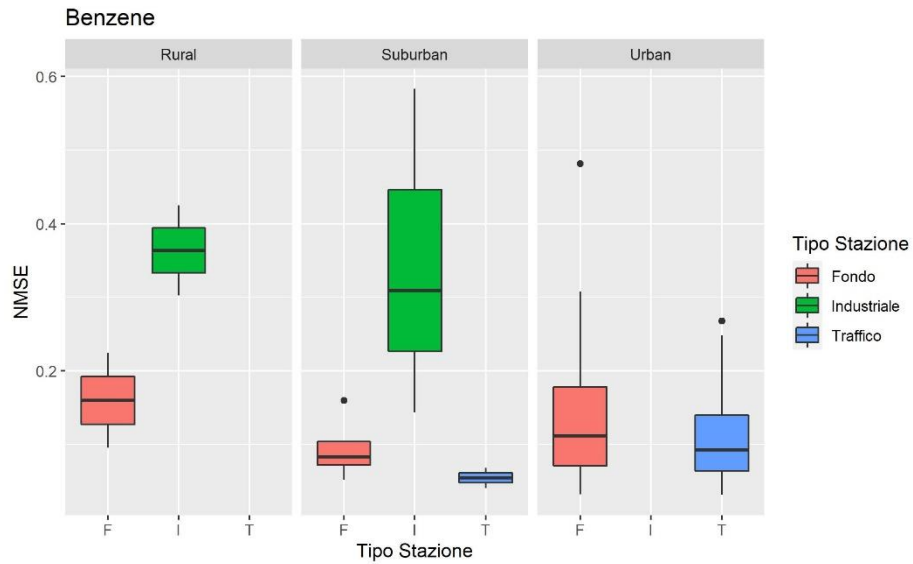


Fig. 2 - Normalized Mean Square Error periodo di validazione (1 gennaio 2020 – 14 febbraio 2020) – Benzene

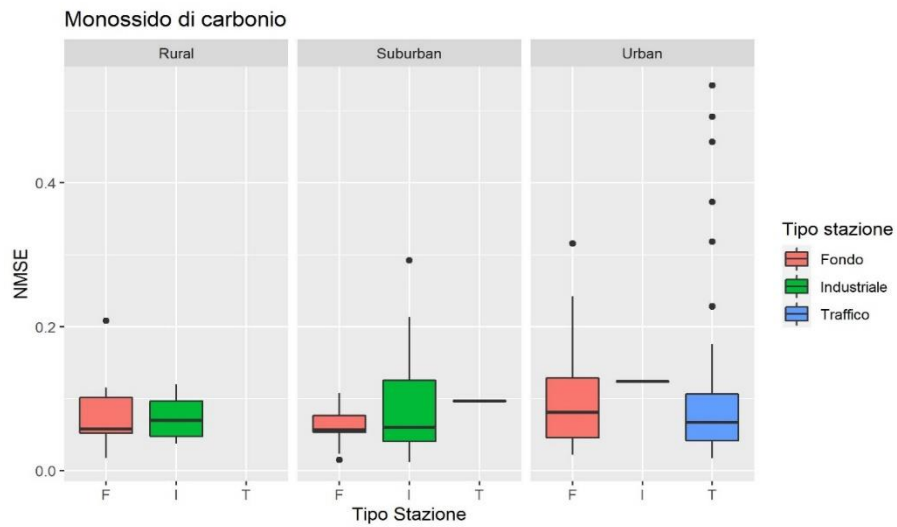


Fig. 3 - Normalized Mean Square Error periodo di validazione (1 gennaio 2020 – 14 febbraio 2020) - Monossido di carbonio

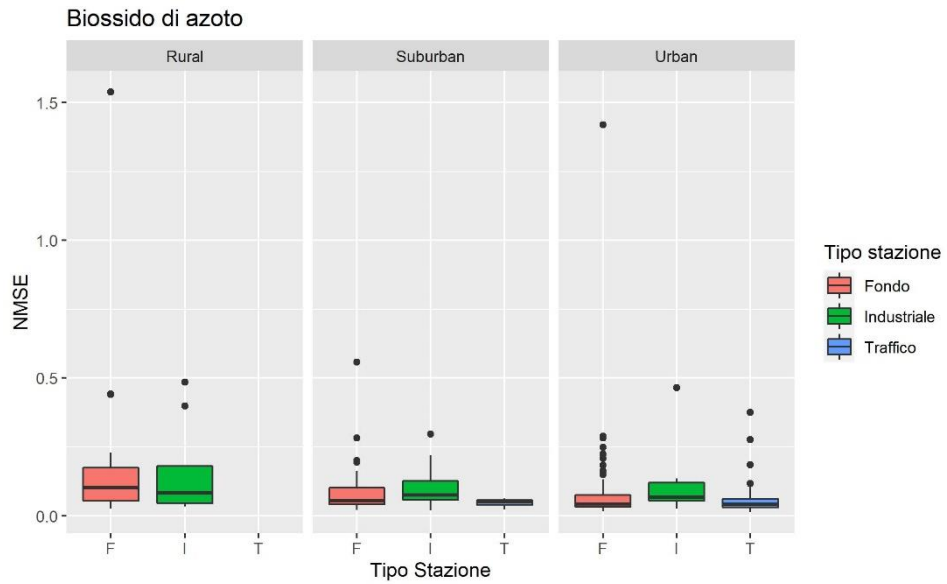


Fig. 4 - Normalized Mean Square Error periodo di validazione (1 gennaio 2020 – 14 febbraio 2020) - Biossido di azoto

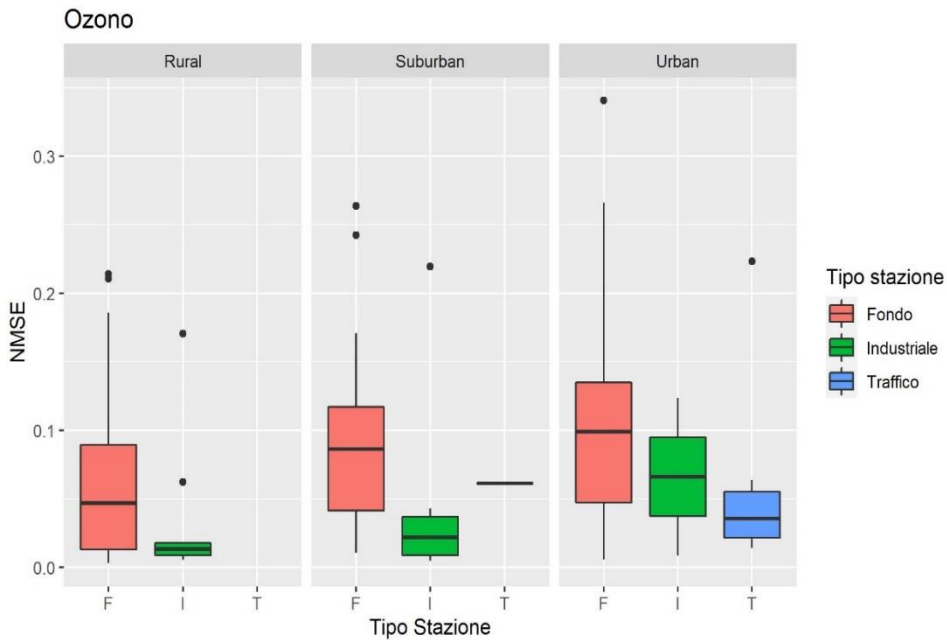


Fig. 5 - Normalized Mean Square Error periodo di validazione (1 gennaio 2020 – 14 febbraio 2020) - Ozono, massimo media mobile 8h

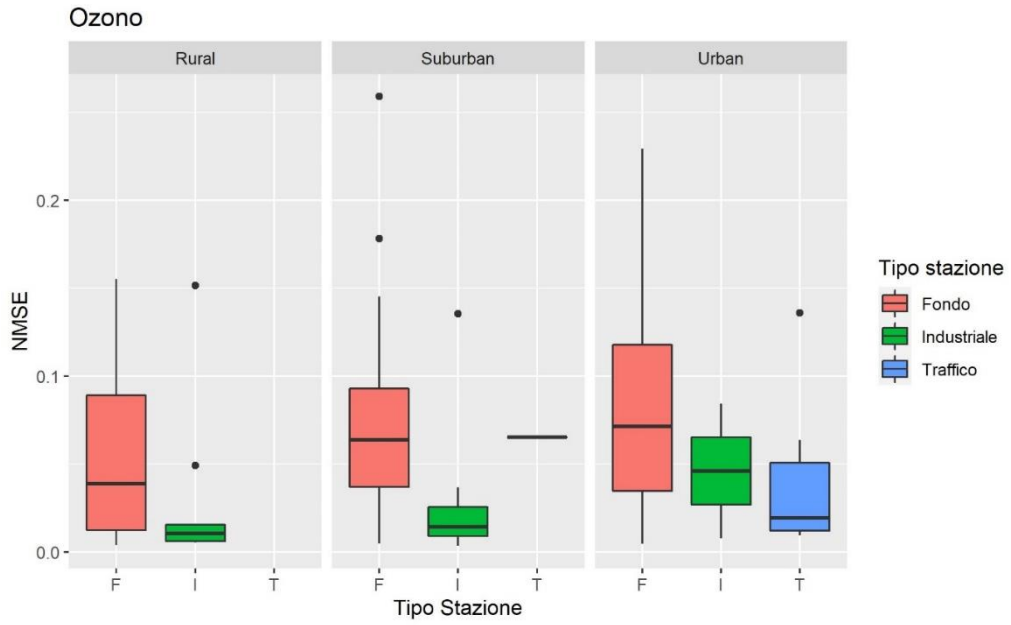


Fig. 6 - Normalized Mean Square Error periodo di validazione (1 gennaio 2020 – 14 febbraio 2020) - Ozono massima media oraria

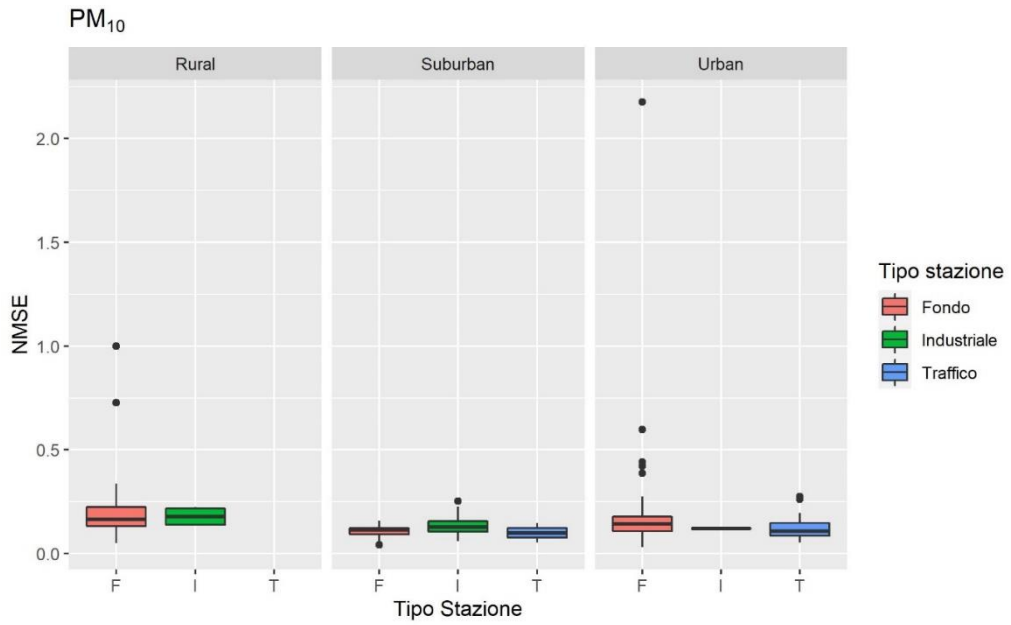


Fig. 7 - Normalized Mean Square Error (1 gennaio 2020 – 14 febbraio 2020) - periodo di validazione PM<sub>10</sub>

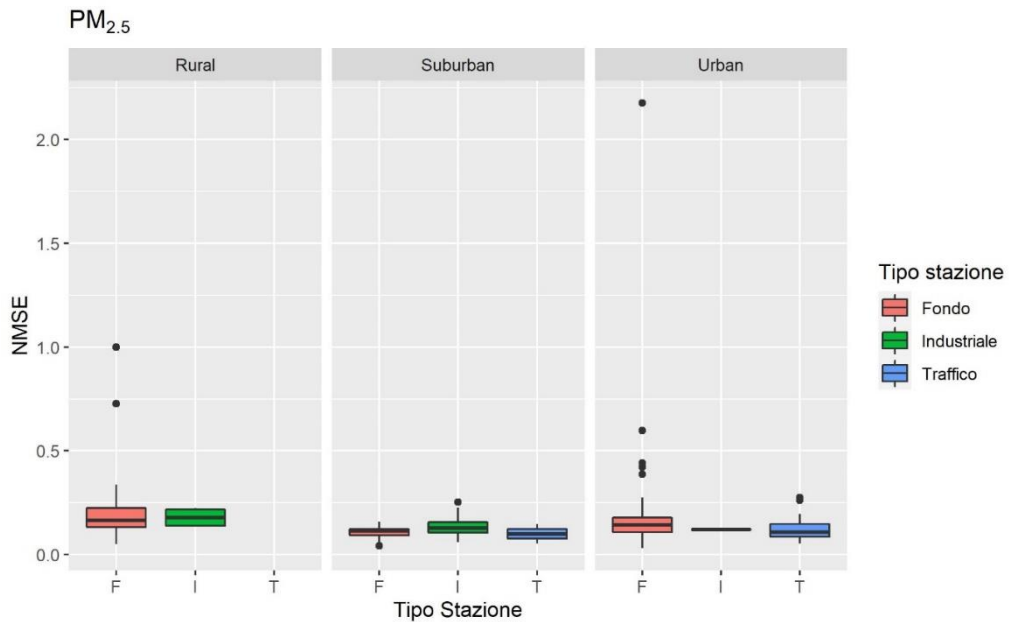


Fig. 8 - Normalized Mean Square Error periodo di validazione (1 gennaio 2020 – 14 febbraio 2020) - PM<sub>2.5</sub>

Le centraline superano la fase di validazione solo se vengono verificati i seguenti criteri:

- $NMSE_{cv} \leq 0.5$
- $NMSE_{t\_oob} \leq 0.5$
- $tr\_r\_squared\_oob \geq 0.5$
- $cv\_r\_squared\_oob \geq 0.5$

dove  $NMSE_{cv}$  è il Normalized Mean Square Error calcolato sul set di validazione,  $NMSE_{t\_oob}$  e  $tr\_r\_squared\_oob$  sono rispettivamente il Normalized Mean Square Error e l' $R^2$  calcolati sul campione out-of-bag del periodo di training (1° gennaio 2013 - 31 dicembre 2019) ed infine  $cv\_r\_squared\_oob$  è l' $R^2$  calcolato sul set di validazione.

Le percentuali di modelli che non hanno superato tutti i criteri di validazione e di valutazione delle performance sono di seguito riportate suddivise per inquinante:

- NO<sub>2</sub>: 2,30%
- PM<sub>2.5</sub>: 15,72%



- PM<sub>10</sub>: 33,93%
- Benzene: 7,09%
- CO: 6.91%
- O<sub>3</sub>: tutti i modelli hanno superato i controlli.

Tali informazioni indicano chiaramente come i modelli sviluppati per l'NO<sub>2</sub> e l'O<sub>3</sub> abbiano prestazioni nettamente migliori rispetto a quelli inerenti agli altri inquinanti e come il particolato, soprattutto quello grossolano, sia risultato più difficile da prevedere rispetto ai gas.

Di seguito sono riportati i grafici che rappresentano la distribuzione delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni. Si evince che, generalmente, le differenze sono prevalentemente negative, il che significa che mediamente questo modello tende a sovrastimare le concentrazioni .

Dalle distribuzioni di densità (grafici a seguire) delle differenze medie nel periodo di validazione per ogni inquinante si nota infatti una leggera asimmetria a sinistra che sta ad indicare una prevalenza delle differenze negative.

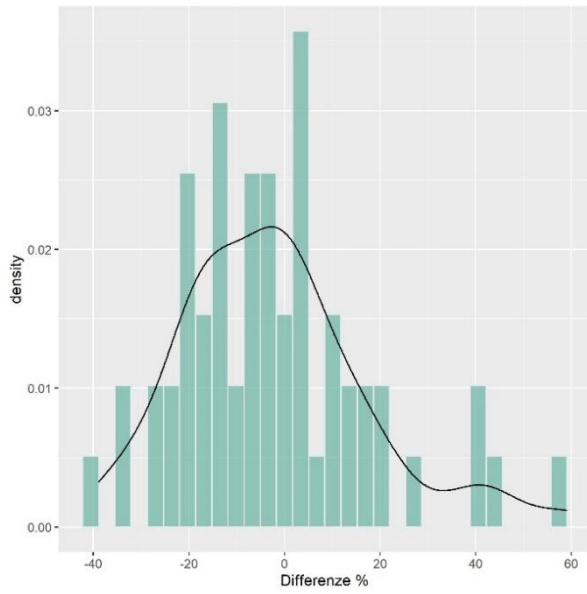


Fig. 9 Distribuzione delle differenze medie nel periodo di validazione Benzene

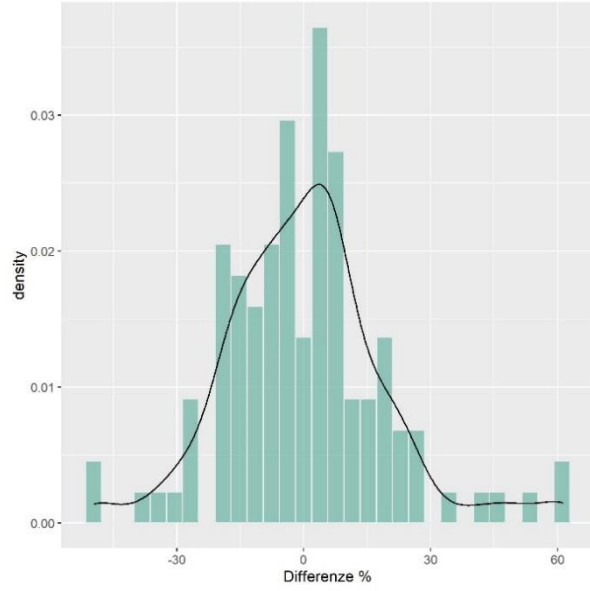


Fig. 10 - Distribuzione delle differenze medie nel periodo di validazione Monossido di carbonio



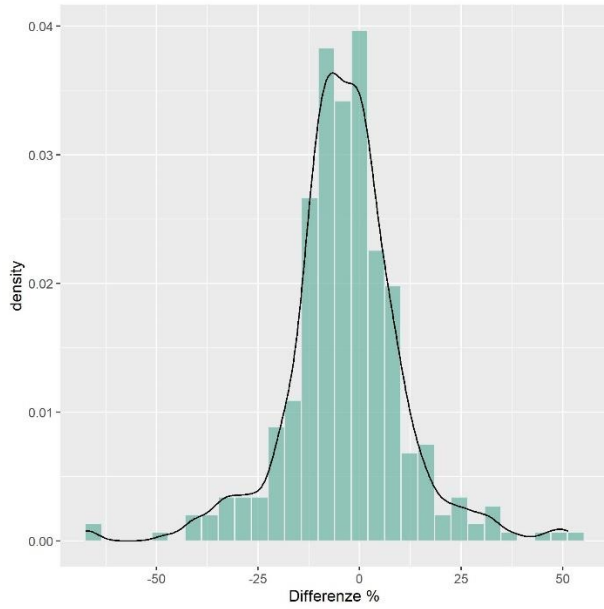


Fig. 11 Distribuzione delle differenze medie nel periodo di validazione - Biossido di azoto

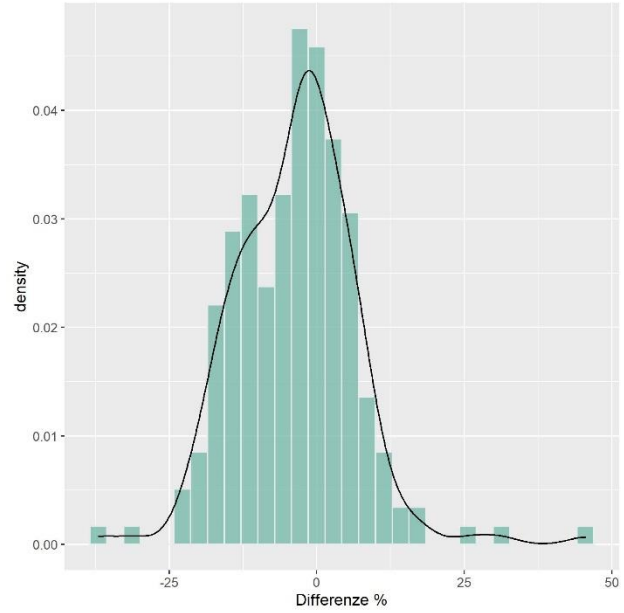


Fig. 12 Distribuzione delle differenze medie nel periodo di validazione Ozono, massimo media oraria

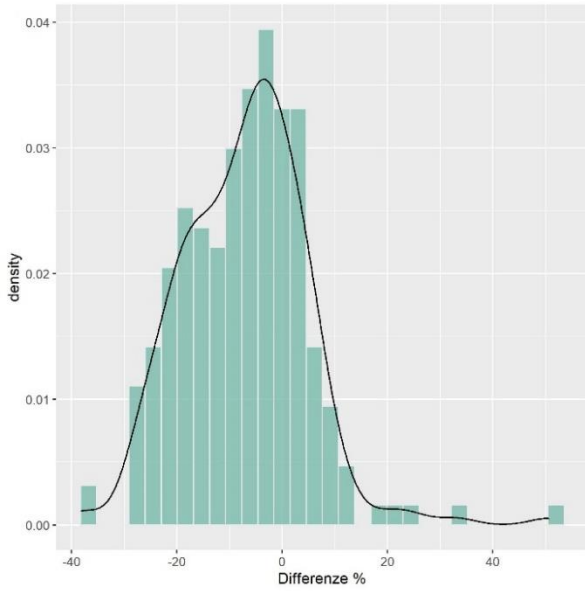


Fig. 13 Distribuzione delle differenze medie nel periodo di validazione Ozono, massimo media mobile 8h

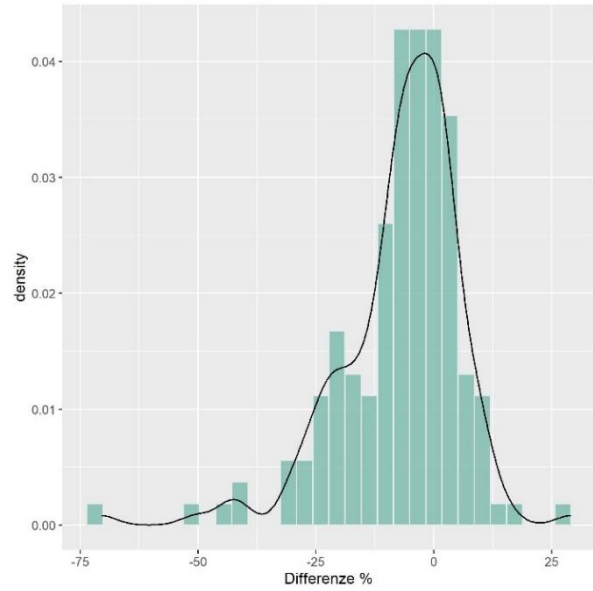


Fig. 14 Distribuzione delle differenze medie nel periodo di validazione – PM<sub>2.5</sub>

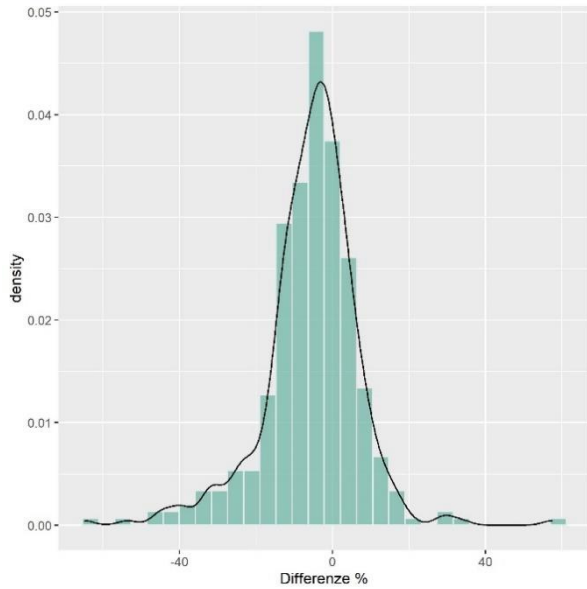


Fig. 15 Distribuzione delle differenze medie nel periodo di validazione – PM<sub>10</sub>

Guardando nel dettaglio i grafici che confrontano i valori osservati e previsti delle singole stazioni e per ogni inquinante sembra che la tendenza del modello sia quella di sovrastimare il dato per valori di concentrazioni bassi e di sottostimare per valori di concentrazione alti; questo comportamento è abbastanza comune in tutti gli algoritmi di ML che tendono a cogliere bene soprattutto i valori modali.

Analizziamo le differenze percentuali più in dettaglio: i box plot riportati di seguito rappresentano le differenze percentuali tra dati osservati e stimati nel periodo di validazione, suddivise per inquinante e tipologia di stazione (background – B, industriale – I, traffico – T; zona rurale – R).

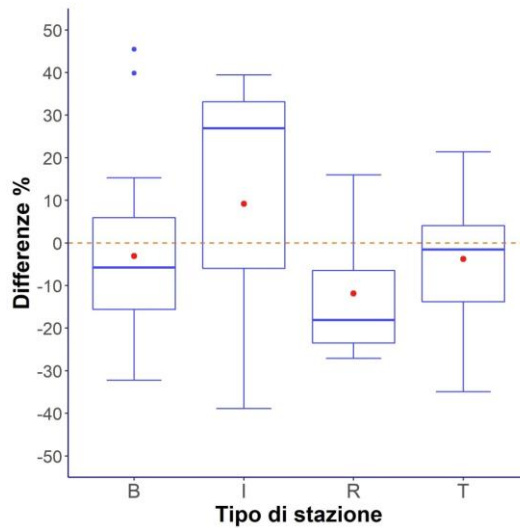


Fig. 16 - Benzene. Box plot delle differenze percentuali, per stazione e tipologia di stazione, tra concentrazione media osservata e concentrazione media stimata dal modello BAU (periodo 1 gennaio 2020 – 14 febbraio 2020).

Per il benzene (Fig. 16), la maggior parte delle stazioni di background e da traffico e la quasi totalità delle stazioni rurali presentano differenze tra dato osservato e previsto negative, quindi c'è una tendenza del modello a sovrastimare le concentrazioni. Considerando il valore assoluto, sia la media che la mediana delle differenze percentuali sono inferiori al 10% per le stazioni di background e da traffico, mentre sono comprese tra il 10 e 20% per le stazioni in zona rurale.

Per quanto riguarda le stazioni di tipo industriale si osserva invece una netta prevalenza di differenze percentuali di segno positivo, con valori previsti mediamente inferiori ai valori osservati. Per quest'ultima tipologia di stazioni le differenze percentuali sono anche di maggiore entità, arrivando a superare il 30% in più di un quarto dei casi.

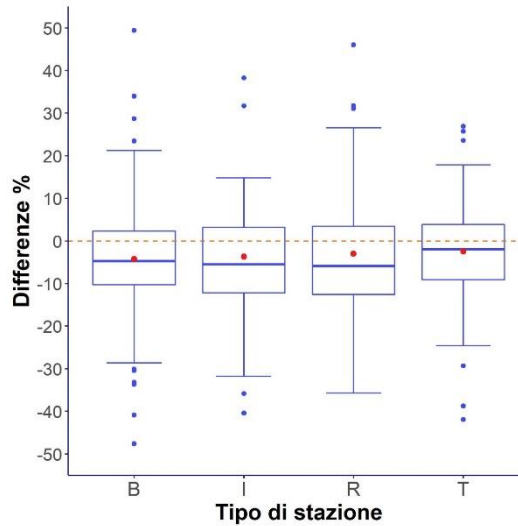


Fig. 17 - Biossido di azoto. Box plot delle differenze percentuali, per stazione e tipologia di stazione, tra concentrazione media osservata e concentrazione media stimata dal modello BAU (periodo 1 gennaio 2020 – 14 febbraio 2020).

Per il biossido di azoto (Fig. 17) la differenza media percentuale è inferiore al 10% (valore assoluto) per la maggior parte delle stazioni, per le 4 tipologie. Media e mediana delle differenze sono negative, indicando che in media il modello tende ad una leggera sovrastima per la maggior parte delle stazioni.

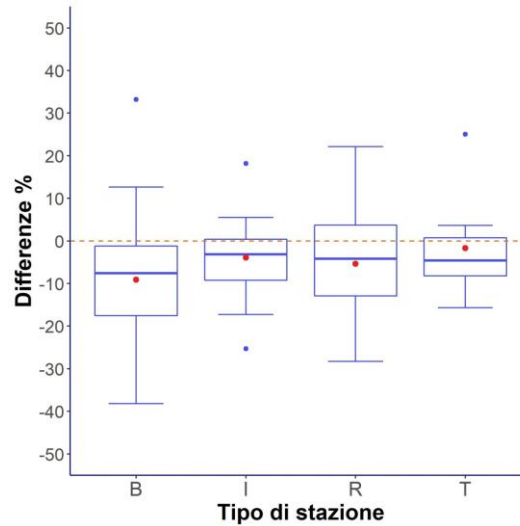


Fig. 18 - Ozono, massimo media mobile 8h. Box plot delle differenze percentuali, per stazione e tipologia di stazione, tra concentrazione media osservata e concentrazione media stimata dal modello BAU (periodo 1 gennaio 2020 – 14 febbraio 2020).

Per l'ozono (massimo della media mobile di 8h), la maggior parte delle stazioni presenta valori previsti mediamente sovrastimati rispetto a quelli osservati, con differenze percentuali che si assestano entro -10% per la maggior parte delle stazioni.

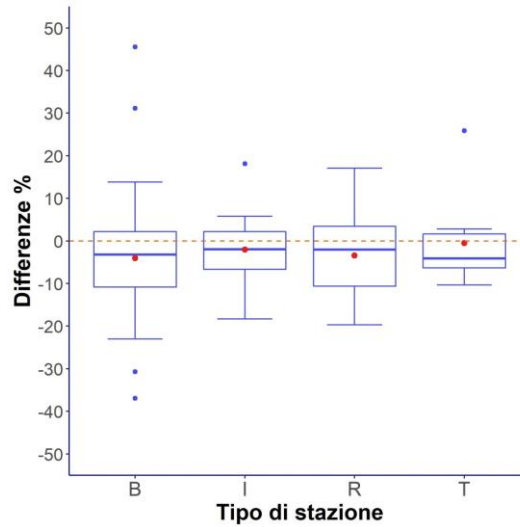


Fig. 19 - Ozono, massima media oraria. Box plot delle differenze percentuali, per stazione e tipologia di stazione, tra concentrazione media osservata e concentrazione media stimata dal modello BAU (periodo 1 gennaio 2020 – 14 febbraio 2020).

Per quel che riguarda la massima media oraria, le differenze tra valori previsti e osservati si comportano all'incirca come quelle osservate per la media mobile su 8 ore, con differenze medie prevalentemente negative e entro -10%, ma con un maggior numero di stazioni che presentano differenze medie positive. Questo risulta coerente con il fatto che questo modello tende a sovrastimare le basse concentrazioni e sottostimare quelle elevate.

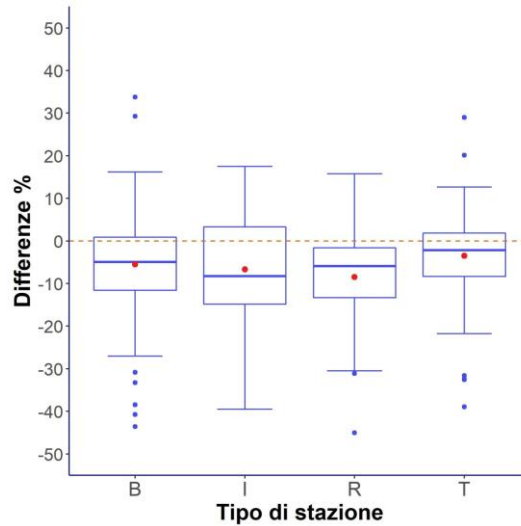


Fig. 20 – PM<sub>10</sub>. Box plot delle differenze percentuali, per stazione e tipologia di stazione, tra concentrazione media osservata e concentrazione media stimata dal modello BAU (periodo 1 gennaio 2020 – 14 febbraio 2020).

Per il PM<sub>10</sub> e il il PM<sub>2.5</sub>, media e mediana delle differenze percentuali sono inferiori al 10% (valore assoluto), per le 4 tipologie di stazioni. Anche in questo caso le differenze risultano prevalentemente di segno negativo, evidenziando che per la maggior parte delle stazioni la media dei valori osservati è inferiore alla media dei valori previsti dal modello.



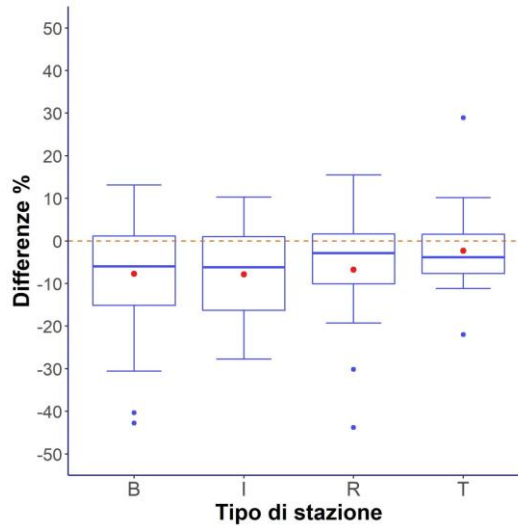


Fig. 21 – PM<sub>2.5</sub>. Box plot delle differenze percentuali, per stazione e tipologia di stazione, tra concentrazione media osservata e concentrazione media stimata dal modello BAU (periodo 1 gennaio 2020 – 14 febbraio 2020).

### 3.2 Valutazione dell'incertezza

Nei grafici seguenti, per ciascun inquinante, sono riportati i box plot delle distribuzioni dello Standard error per tutto il periodo di lockdown calcolato seguendo il metodo Jackknife-after-Bootstrap (Wager, Hastie & Efron [2014]) come più dettagliatamente spiegato nel paragrafo 2.3.2. I dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.

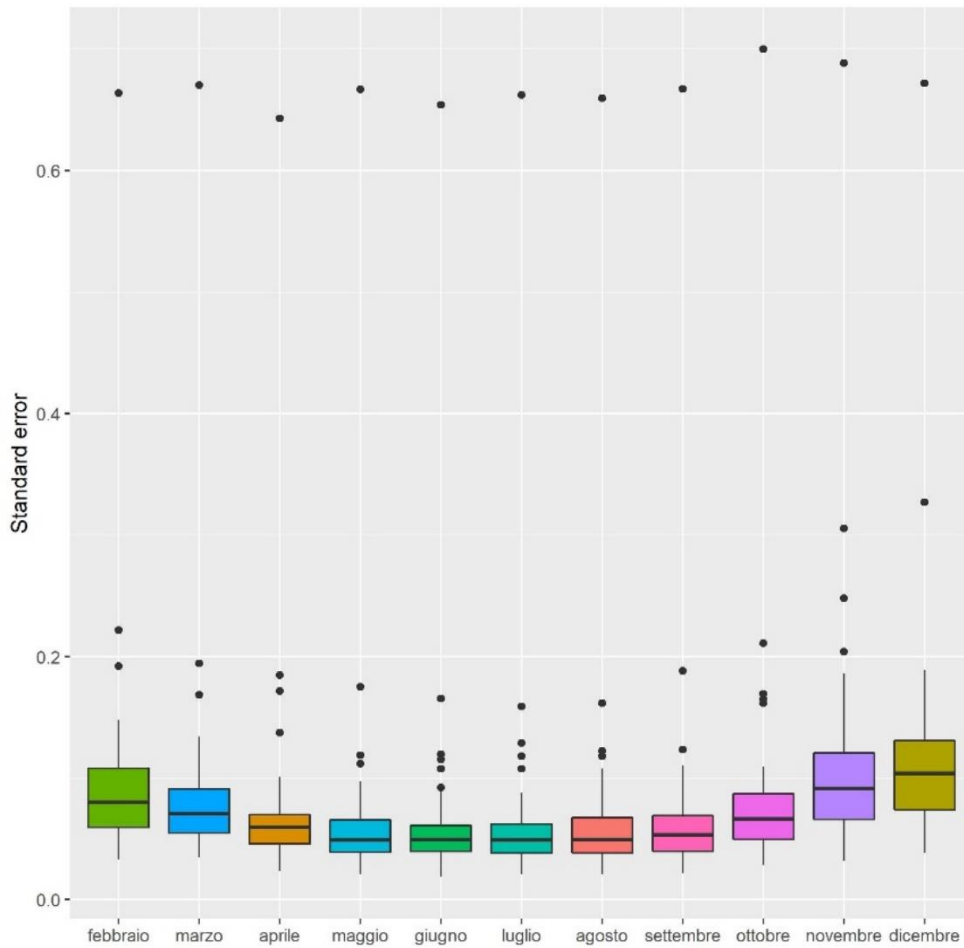


Fig. 22 Benzene Distribuzione dello Standard error delle previsioni ( $\mu\text{g}/\text{m}^3$ ) per mese

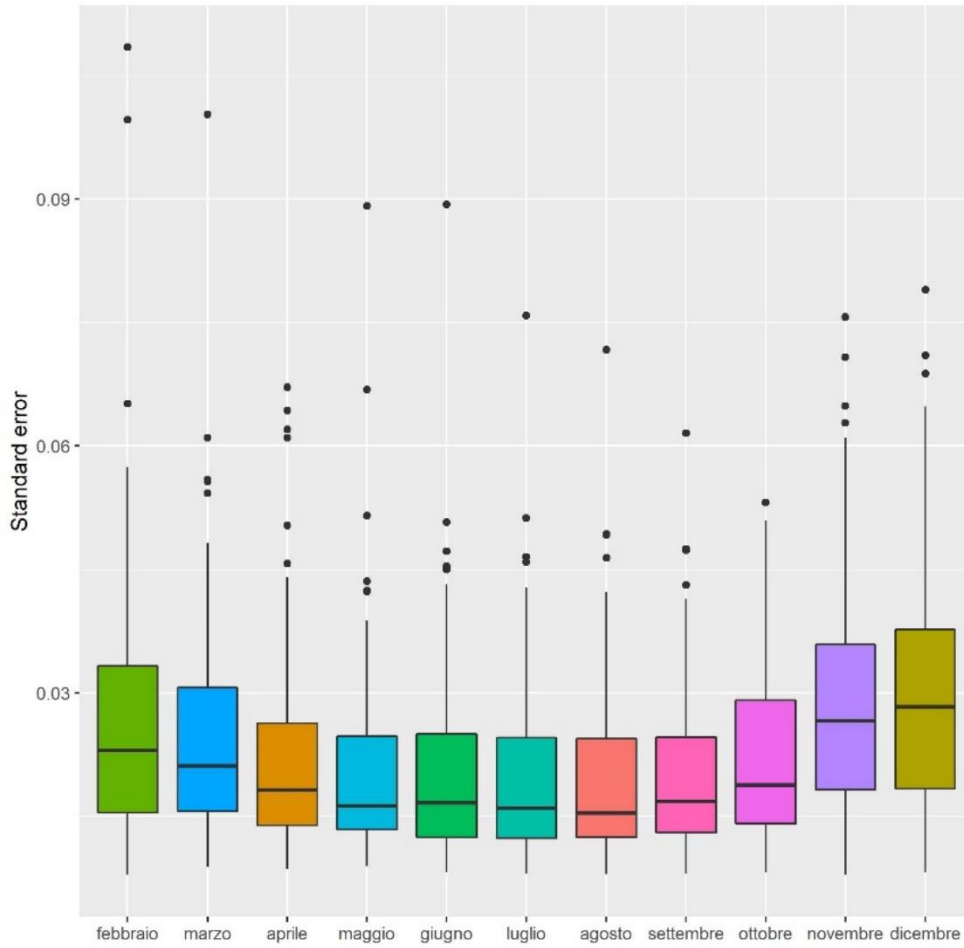


Fig. 23 Monossido di Carbonio Distribuzione dello Standard error delle previsioni ( $\mu\text{g}/\text{m}^3$ ) per mese

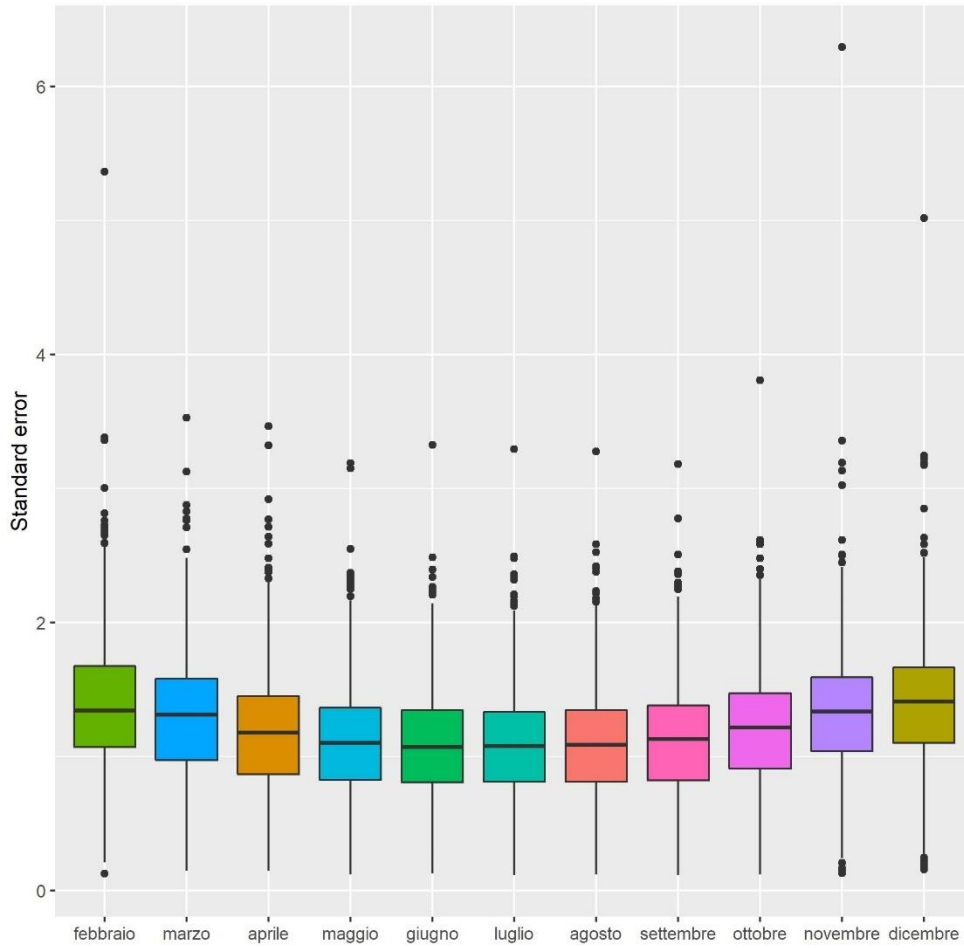


Fig. 24 Biossido di azoto Distribuzione dello Standard error delle previsioni ( $\mu\text{g}/\text{m}^3$ ) per mese

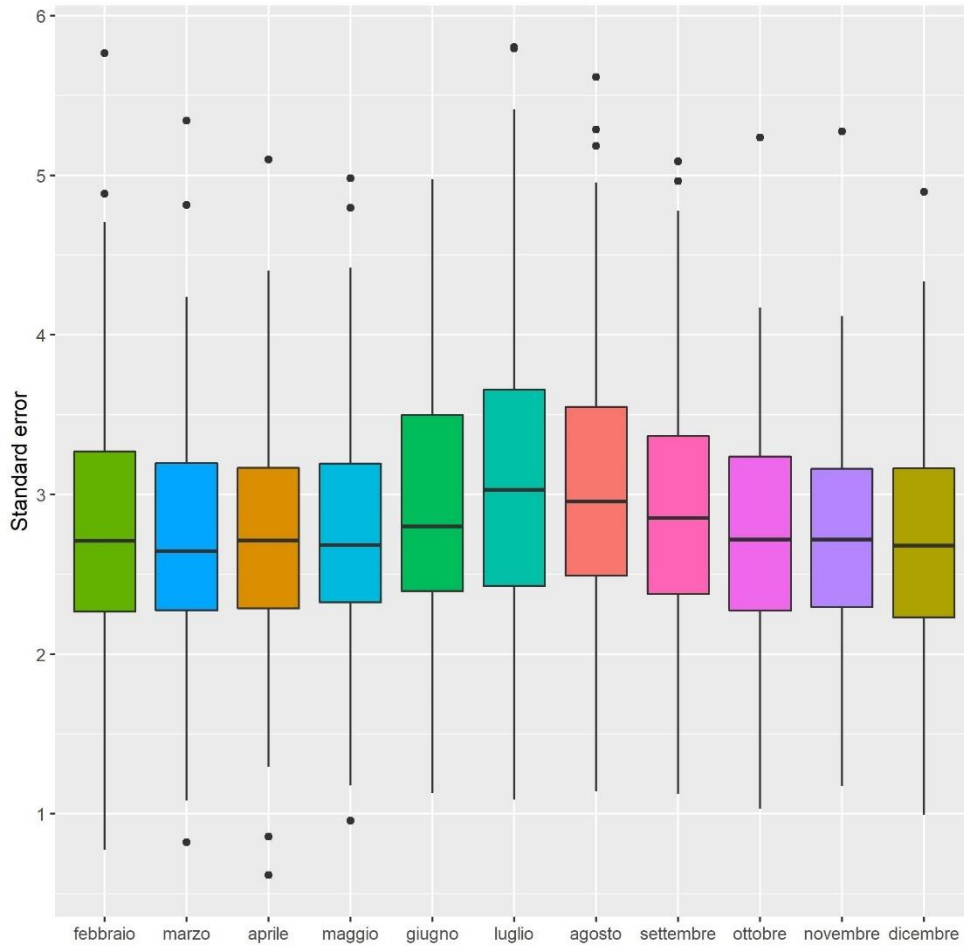


Fig. 25 Ozono massima media oraria Distribuzione dello Standard error delle previsioni ( $\mu\text{g}/\text{m}^3$ ) per mese

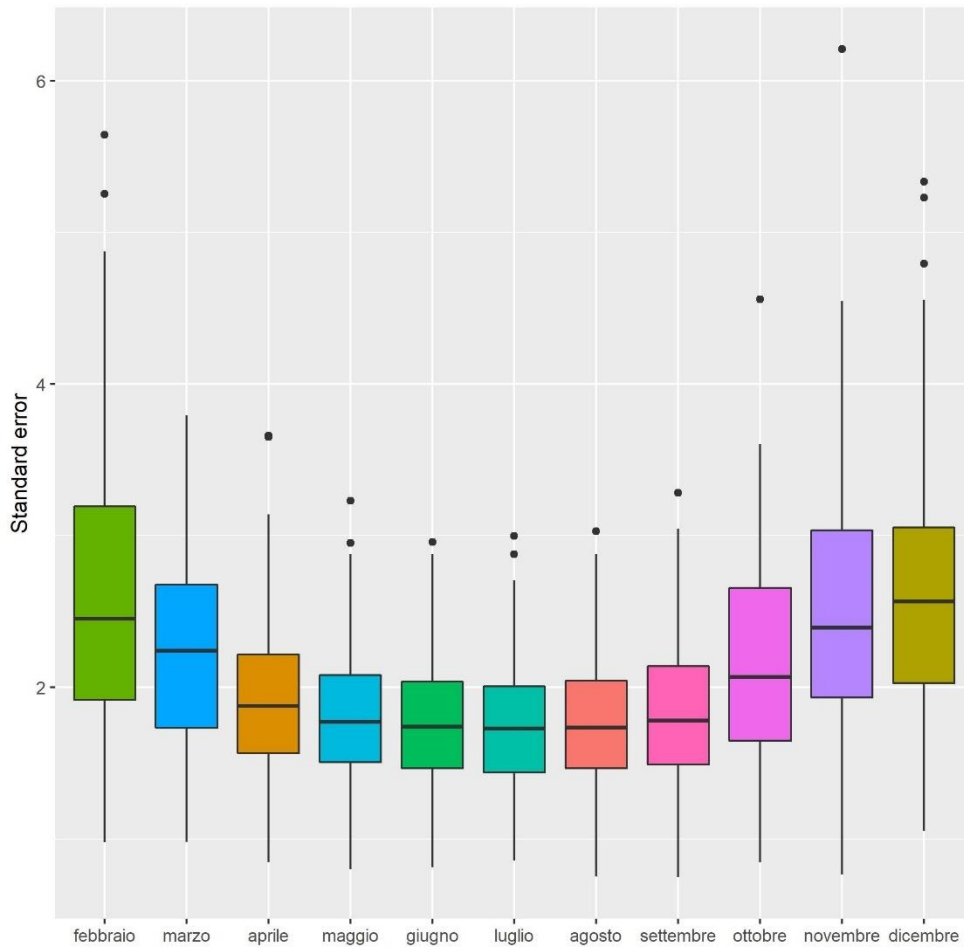


Fig. 26 PM<sub>10</sub> Distribuzione dello Standard error delle previsioni ( $\mu\text{g}/\text{m}^3$ ) per mese

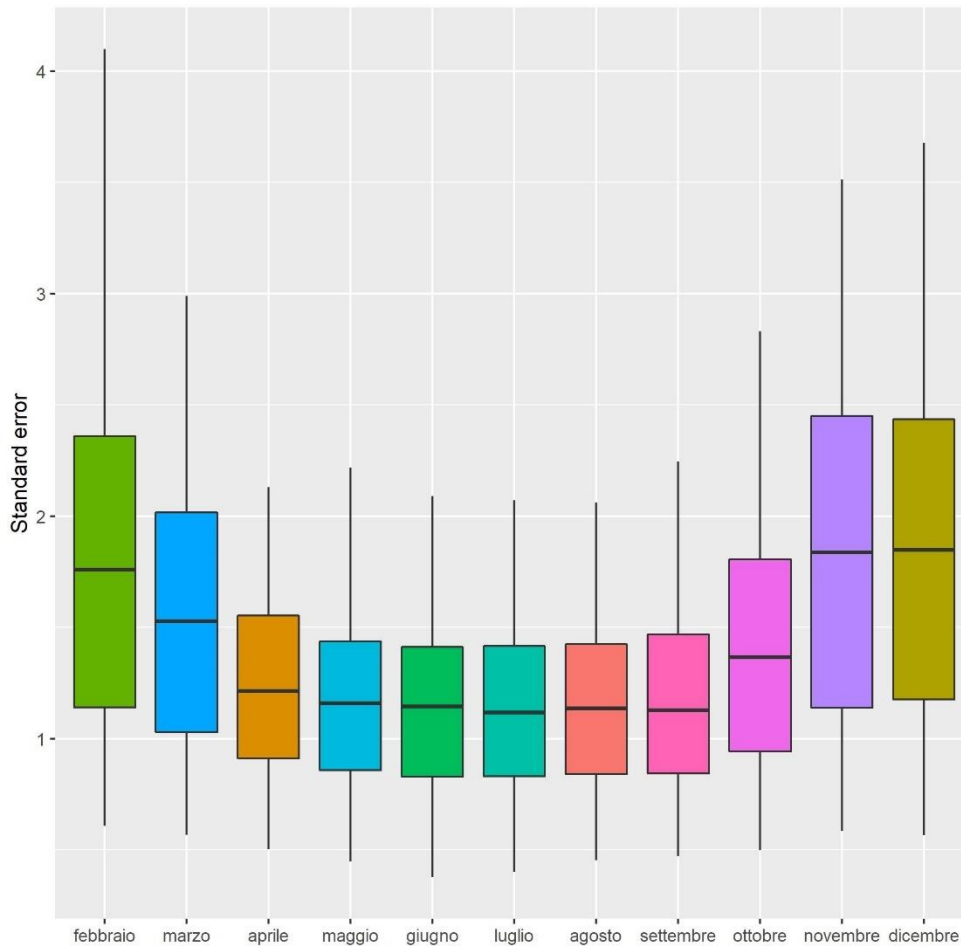


Fig. 27 PM<sub>2.5</sub> Distribuzione dello Standard error delle previsioni ( $\mu\text{g}/\text{m}^3$ ) per mese

### 3.3 Differenze dell'anno 2020 rispetto all'atteso

Per valutare l'effetto del lockdown sulle concentrazioni degli inquinanti, si sono osservate le differenze tra il dato previsto dal modello e le concentrazioni osservate. Dal momento che il modello è stato addestrato sugli anni precedenti il periodo di pandemia (2013 – 2019), le concentrazioni previste sono quelle che si sarebbero osservate in assenza di misure restrittive, con le condizioni meteorologiche che si sono verificate nei mesi in studio. La differenza tra i dati - a meno



dell'incertezza di stima del modello - può quindi ragionevolmente essere imputabile all'effetto delle limitazioni introdotte a seguito della pandemia. È evidente tuttavia che non è possibile con questo metodo valutare l'effetto sulla qualità dell'aria della riduzione delle emissioni dei singoli comparti emissivi e soprattutto cogliere l'importanza relativa di questi; occorre anche tener conto della possibilità che le emissioni siano diminuite per alcuni comparti e aumentate o rimaste pressoché invariate in altri. Per un approfondito esame di questi aspetti si rimanda ai report dell'obiettivo 2.

Di seguito si riportano i dettagli per singolo inquinante: tutti i dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.

### 3.3.1 Benzene

Nelle cartine seguenti sono riportate le differenze medie percentuali tra il dato osservato e il dato previsto dal modello per le singole stazioni su cui è stato rilevato il benzene. Le tre mappe si riferiscono ai periodi 1 marzo – 30 aprile, 1 maggio – 30 settembre e 1 ottobre - 31 dicembre 2020.





Fig. 28 Benzene Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)



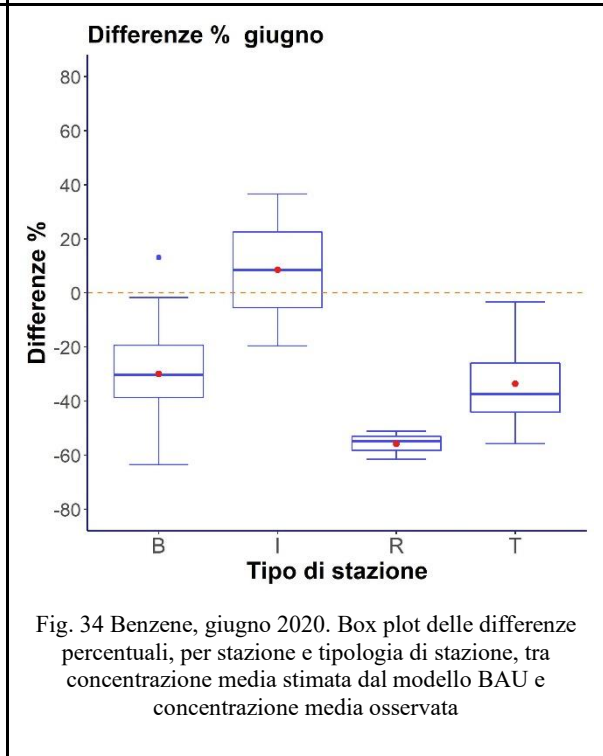
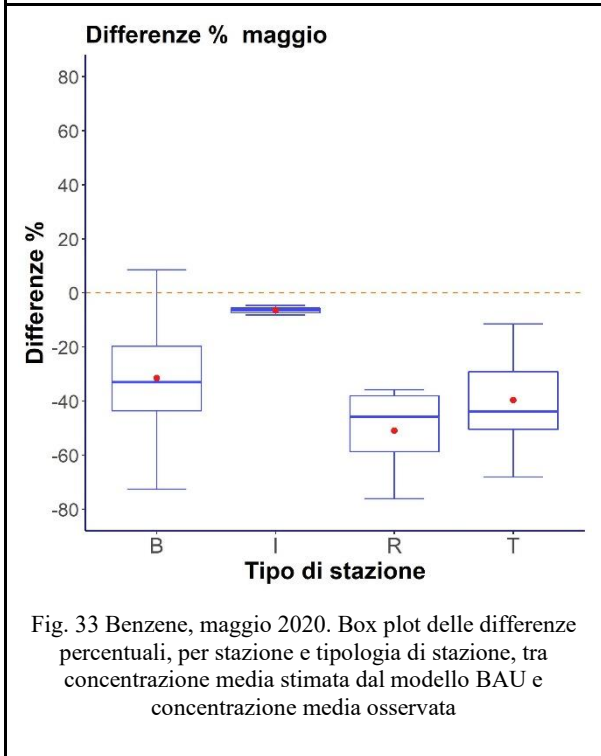
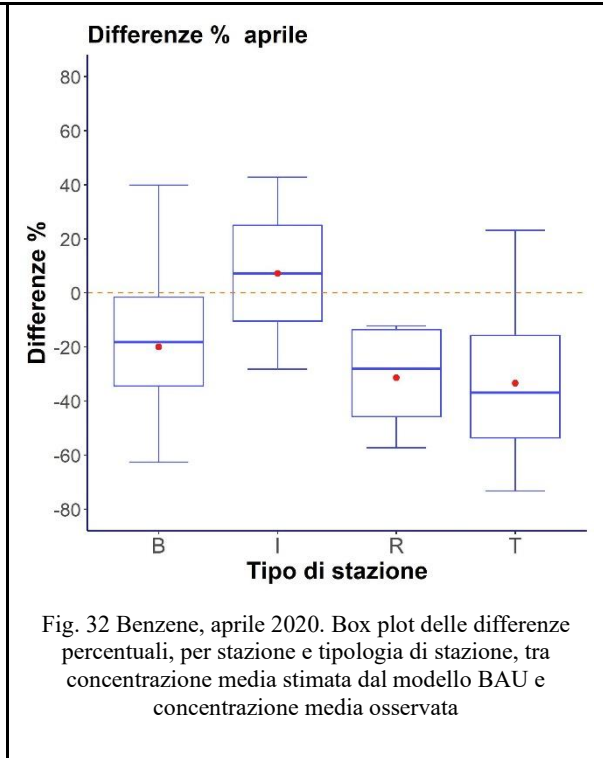
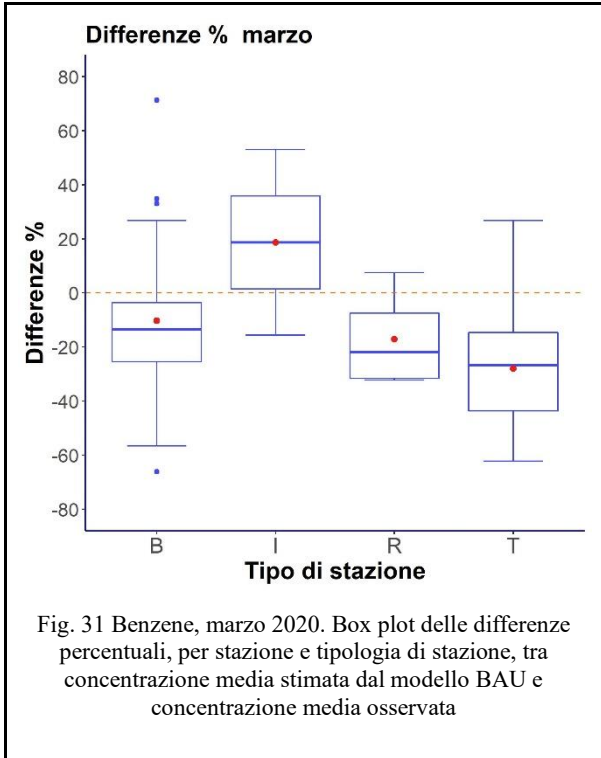
Fig. 29 Benzene Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)

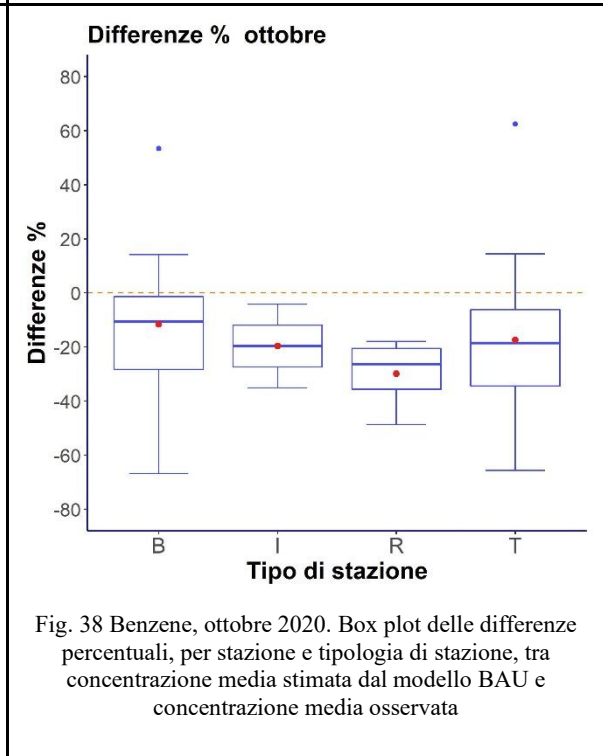
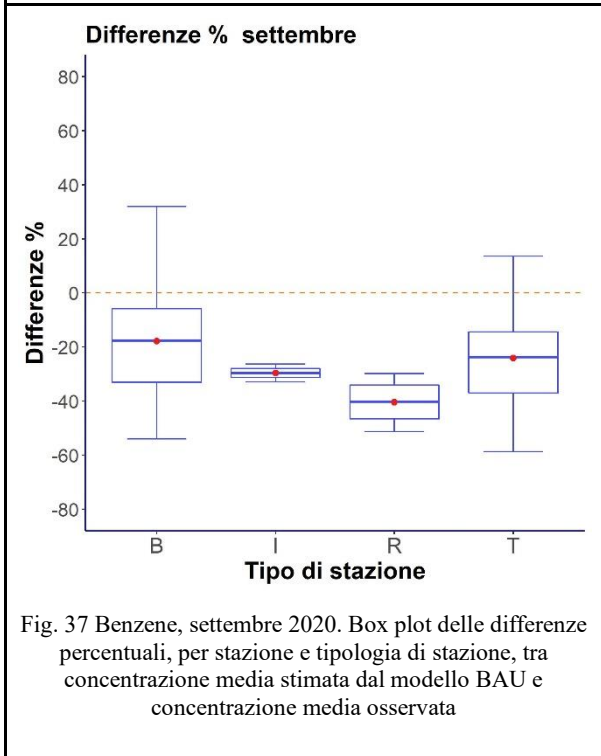
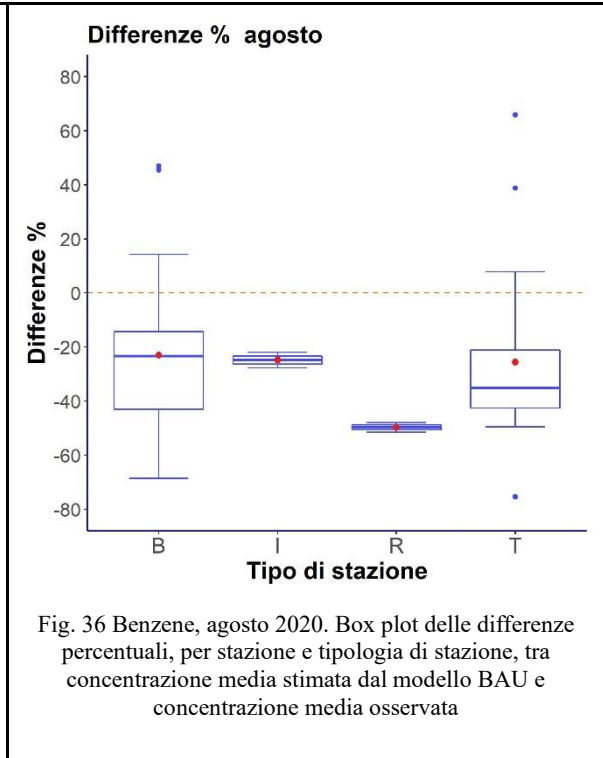
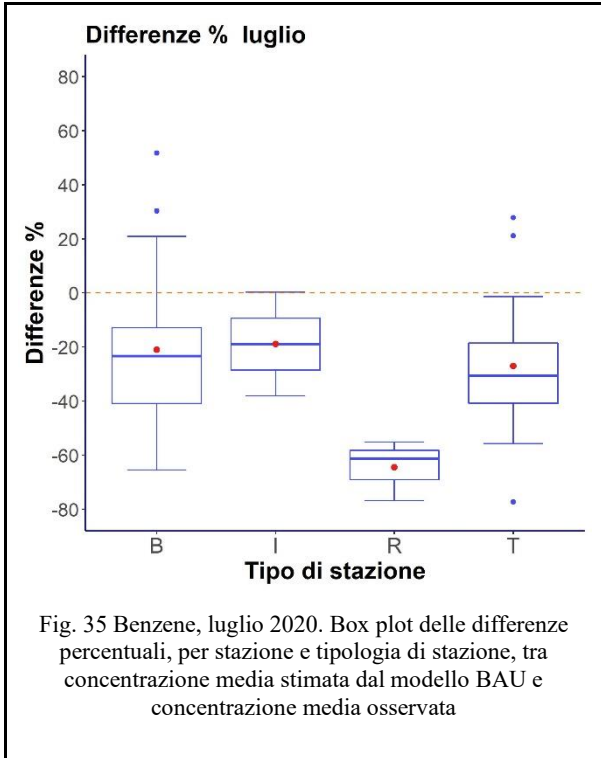


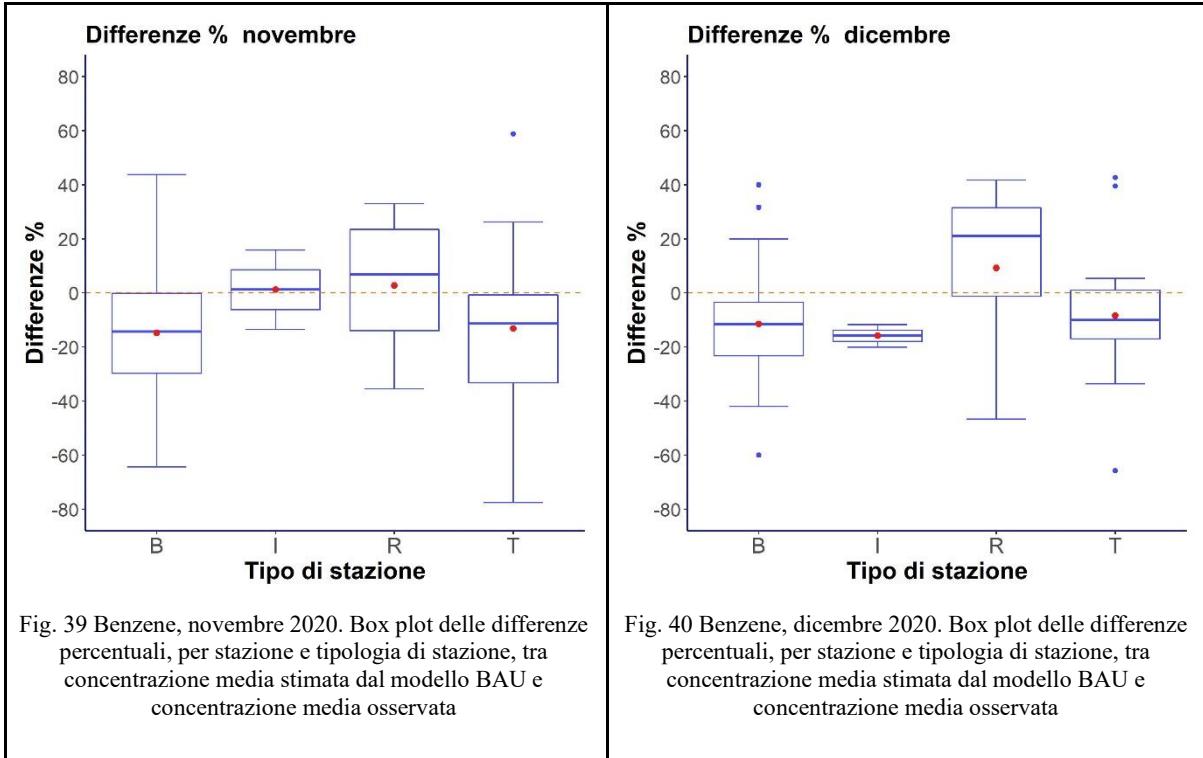
Fig. 30 Benzene Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)

Non si evidenzia un gradiente Nord-Sud che faccia ipotizzare un diverso comportamento delle stazioni in base alla latitudine. La quasi totalità delle stazioni presenta differenze negative, con valori osservati inferiori ai valori previsti dal modello. Questo dato era atteso sia per il fatto che durante il periodo in studio sono state messe in campo restrizioni agli spostamenti a causa della pandemia di coronavirus, sia per il fatto che il modello tende ad effettuare previsioni che mediamente sovrastimano il dato osservato, come discusso nel paragrafo 3.1.

I boxplot seguenti rappresentano la distribuzione, mese per mese, delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni, raggruppate per tipologia di stazione.







Le differenze percentuali risultano generalmente più basse di quelle osservate durante il periodo di validazione per tutto l'anno e non solo per il periodo di lockdown di marzo-maggio. Questo suggerisce che le limitazioni agli spostamenti indotte dal contenimento della pandemia abbiano prodotto una diminuzione delle concentrazioni di tale inquinante, anche nei mesi successivi al lock-down primaverile.

Le stazioni da traffico, coerentemente con quanto atteso, mostrano per tutti i mesi differenze percentuali medie negative per quasi tutte le stazioni, con una differenza media inferiore a quelle riscontrata durante il periodo di validazione e quindi imputabile alle restrizioni.

Le stazioni rurali presentano un calo importante durante molti mesi dell'anno e in particolar modo durante i mesi estivi: da maggio a settembre la media delle differenze risulta inferiore a -40%, mentre durante il periodo di validazione era di poco inferiore a -10%. Le differenze sono negative e in valore assoluto maggiori rispetto a quelle osservate durante il periodo di validazione anche nei mesi di marzo, aprile e ottobre, mentre nei mesi di novembre e dicembre media e mediana

delle differenze risultano lievemente positive, ad indicare che in questi mesi non vi sia stato un calo di questo inquinante in questa tipologia di stazioni.

Le stazioni industriali (sicuramente in numero ridotto) presentano un andamento piuttosto variabile durante l'anno, con differenze percentuali inaspettatamente positive nei mesi di marzo, aprile, giugno e agosto (così si era osservato anche durante il periodo di validazione) e negative o prossime allo zero nei restanti mesi.

Le stazioni di background presentano concentrazioni negative in quasi tutte le stazioni, per tutti i mesi indagati. Le differenze medie risultano sempre maggiori, in valore assoluto, rispetto a quelle osservate durante il periodo di validazione, indicando che per questa tipologia di stazione si è osservato un calo, relativamente a questo inquinante, durante tutti i mesi successivi a febbraio 2020. Le maggiori riduzioni si sono osservate nei mesi di maggio e giugno.

Nella tabella seguente sono riportati gli indicatori statistici di base relativi alla variazione assoluta della concentrazione di Benzene durante il 2020 per tipo di stazione e per mese.

Tab. 4 Statistiche descrittive delle variazioni assolute di Benzene, nelle centraline valide, per tipo di stazione e per mese ( $\mu\text{g}/\text{m}^3$ )

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err
<b>Marzo</b>	-0,36	-0,07	-0,14	0,11	0,07	-0,68	-0,22	-0,16	0,2	0,22	-0,58	-0,36	-0,3	-0,09	0,08
<b>Aprile</b>	-0,27	-0,12	-0,13	0	0,06	-0,73	-0,38	-0,45	0,07	0,2	-0,66	-0,39	-0,27	-0,1	0,07
<b>Maggio</b>	-0,27	-0,17	-0,15	-0,07	0,05	-0,63	-0,25	-0,35	-0,14	0,2	-0,48	-0,33	-0,23	-0,12	0,06
<b>Giugno</b>	-0,24	-0,15	-0,12	-0,06	0,05	-0,65	-0,03	-0,46	-0,13	0,2	-0,37	-0,25	-0,18	-0,09	0,06
<b>Luglio</b>	-0,22	-0,12	-0,12	-0,05	0,05	-0,84	-0,61	-0,71	-0,19	0,2	-0,3	-0,16	-0,13	-0,06	0,06
<b>Agosto</b>	-0,24	-0,14	-0,13	-0,04	0,05	-0,76	-0,49	-0,46	-0,2	0,21	-0,37	-0,18	-0,15	-0,05	0,06
<b>Settembre</b>	-0,23	-0,12	-0,1	0	0,05	-0,37	-0,3	-0,22	-0,14	0,22	-0,33	-0,21	-0,16	-0,04	0,07
<b>Ottobre</b>	-0,28	-0,12	-0,11	0,04	0,07	-0,27	-0,18	-0,11	-0,05	0,16	-0,43	-0,24	-0,18	0,01	0,08
<b>Novembre</b>	-0,52	-0,27	-0,22	0,08	0,1	-0,21	-0,03	0	0,15	0,2	-0,76	-0,33	-0,2	0,1	0,11
<b>Dicembre</b>	-0,59	-0,25	-0,24	0,11	0,11	-0,18	-0,08	-0,03	0,23	0,21	-0,51	-0,09	-0,1	0,25	0,12

Tab. 5 Statistiche descrittive delle variazioni di Benzene, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-32,89	-5,77	-14,15	13,06	-45,24	-5,74	-21,94	32,11	-48,79	-27,27	-31,21	-11,09
Aprile	-42,37	-19,16	-23,90	-0,14	-60,02	-21,93	-39,58	3,99	-57,02	-34,09	-40,27	-19,12
Maggio	-51,97	-32,96	-35,44	-18,48	-66,10	-32,01	-44,17	-17,71	-58,53	-39,35	-45,80	-31,25
Giugno	-49,37	-31,63	-34,22	-15,71	-72,17	-28,96	-51,23	-21,20	-51,90	-34,28	-38,84	-22,34
Luglio	-48,51	-23,32	-32,14	-12,33	-84,07	-50,10	-67,68	-36,90	-47,95	-26,39	-32,89	-16,06
Agosto	-49,70	-27,20	-34,20	-11,32	-77,08	-45,14	-48,82	-25,11	-50,77	-25,55	-34,71	-16,31
Settembre	-42,13	-18,79	-23,08	0,94	-46,77	-32,44	-39,96	-26,11	-45,22	-23,53	-26,31	-6,99
Ottobre	-36,79	-13,32	-17,21	5,29	-44,06	-20,52	-30,82	-11,37	-41,69	-16,78	-21,94	1,19
Novembre	-39,84	-14,12	-17,23	6,81	-27,26	4,26	0,29	27,32	-39,19	-13,21	-16,19	8,73
Dicembre	-39,38	-14,11	-18,49	7,20	-28,94	5,97	-3,89	31,71	-33,07	-3,31	-7,46	16,89

Tab. 6 Significatività delle variazioni di Benzene, nelle centraline valide, per tipo di variazione e per mese (%)

mese	Δ non significativa			Δ significativa		
	Δ negativa	Δ positiva	Totale non significativa	Δ negativa	Δ positiva	Totale significativa
Marzo	13,28%	8,73%	22,01%	59,08%	18,91%	77,99%
Aprile	14,58%	9,44%	24,02%	64,16%	11,82%	75,98%
Maggio	14,18%	4,66%	18,84%	75,99%	5,17%	81,16%
Giugno	21,07%	5,88%	26,96%	68,71%	4,33%	73,04%
Luglio	20,81%	8,70%	29,50%	63,35%	7,14%	70,50%
Agosto	18,61%	9,99%	28,61%	64,21%	7,18%	71,39%
Settembre	20,89%	10,71%	31,60%	57,82%	10,58%	68,40%
Ottobre	18,12%	11,04%	29,16%	54,22%	16,62%	70,84%
Novembre	13,12%	12,18%	25,30%	54,28%	20,41%	74,70%
Dicembre	13,86%	11,39%	25,25%	48,42%	26,33%	74,75%
<b>Totale</b>	<b>16,85%</b>	<b>9,25%</b>	<b>26,11%</b>	<b>61,07%</b>	<b>12,82%</b>	<b>73,89%</b>



### 3.3.2 Monossido di Carbonio

Nelle mappe seguenti sono riportate le differenze medie percentuali tra il dato osservato e previsto dal modello per le singole stazioni su cui è stato rilevato il monossido di carbonio. Le tre mappe si riferiscono ai periodi 1 marzo – 30 aprile, 1 maggio – 30 settembre e 1 ottobre - 31 dicembre 2020. I dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.





Leaflet | © OpenStreetMap contributors, CC-BY-SA, Tiles © Esri — Source: US National Park Service, Map tiles by Sta Design, CC BY 3.0 — Map data © OpenStreetMap contributors

Fig. 41 Monossido di Carbonio Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)



Leaflet | © OpenStreetMap contributors, CC-BY-SA, Tiles © Esri — Source: US National Park Service, Map tiles by Sta Design, CC BY 3.0 — Map data © OpenStreetMap contributors

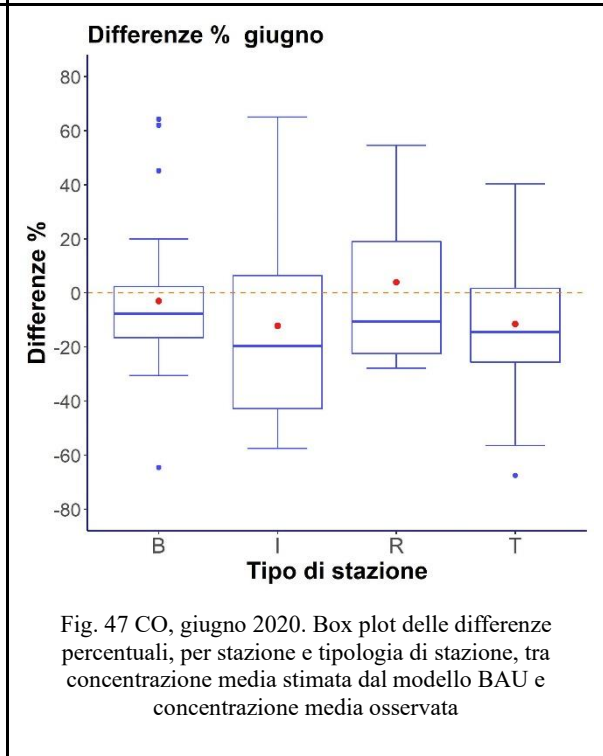
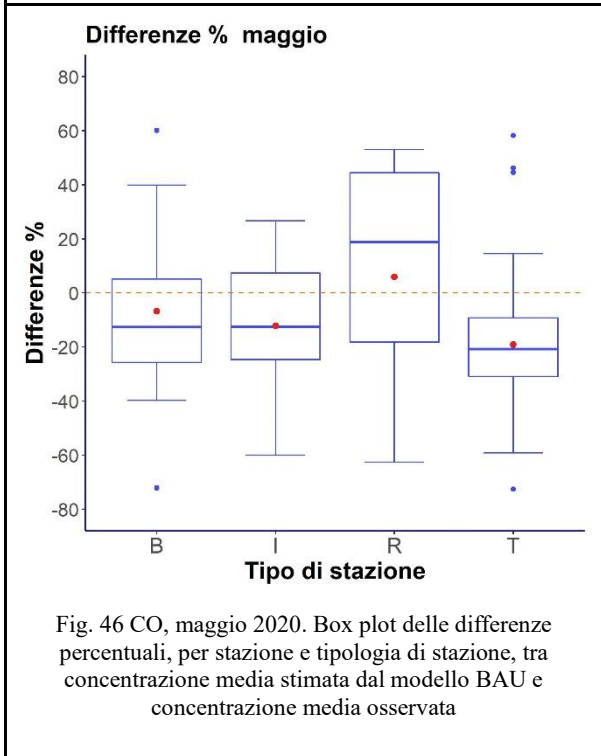
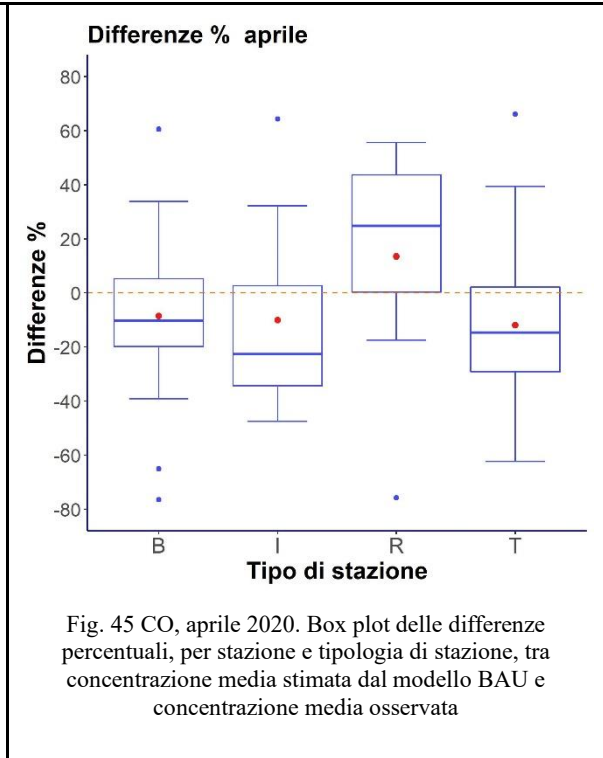
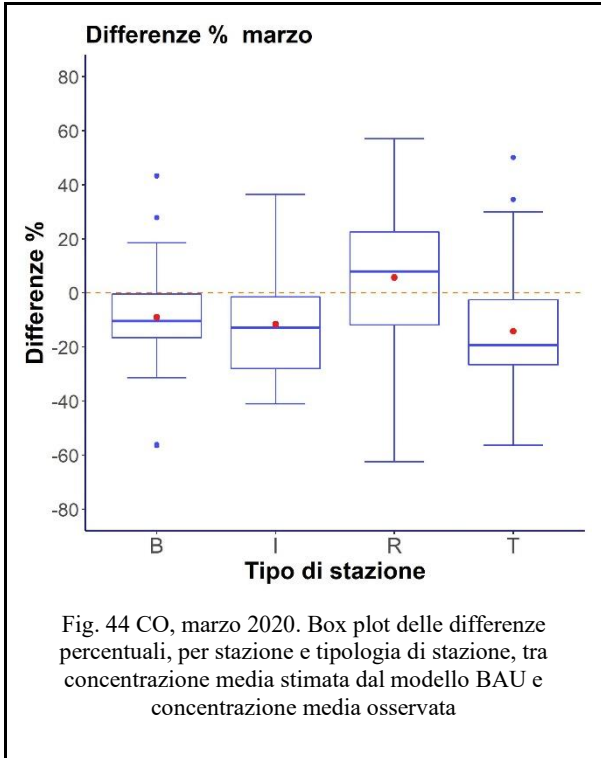
Fig. 42 Monossido di Carbonio Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)

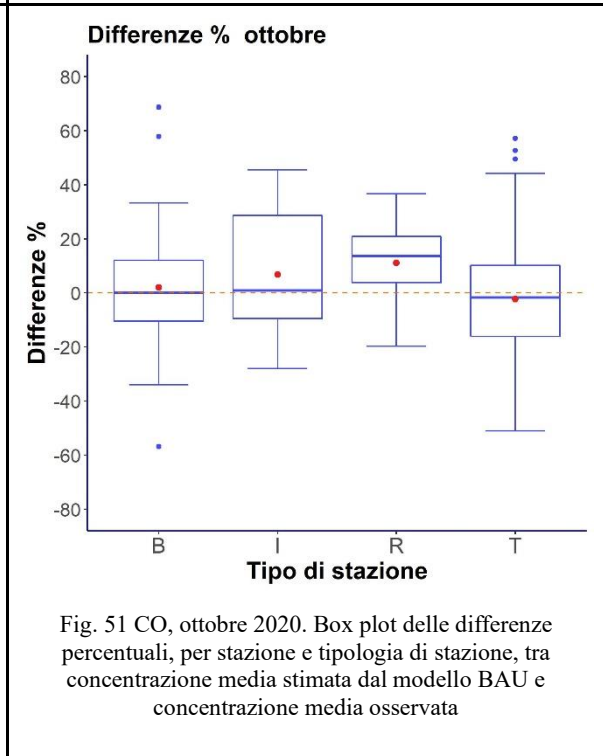
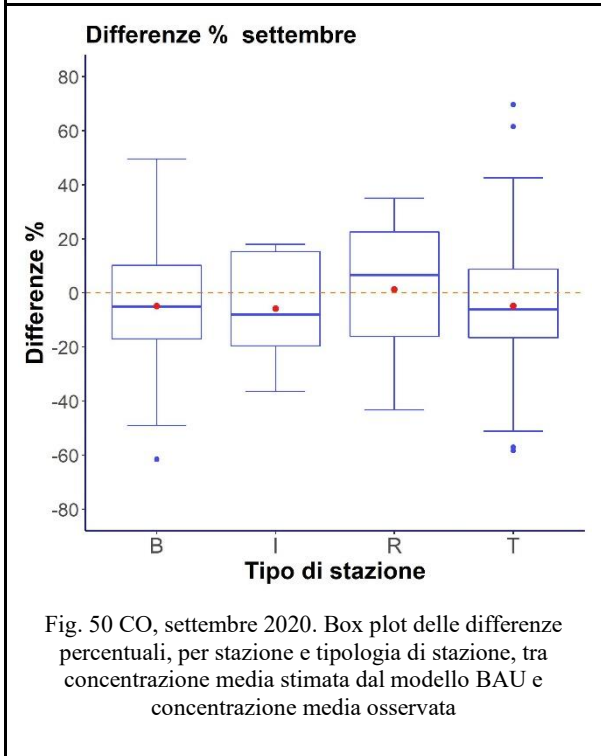
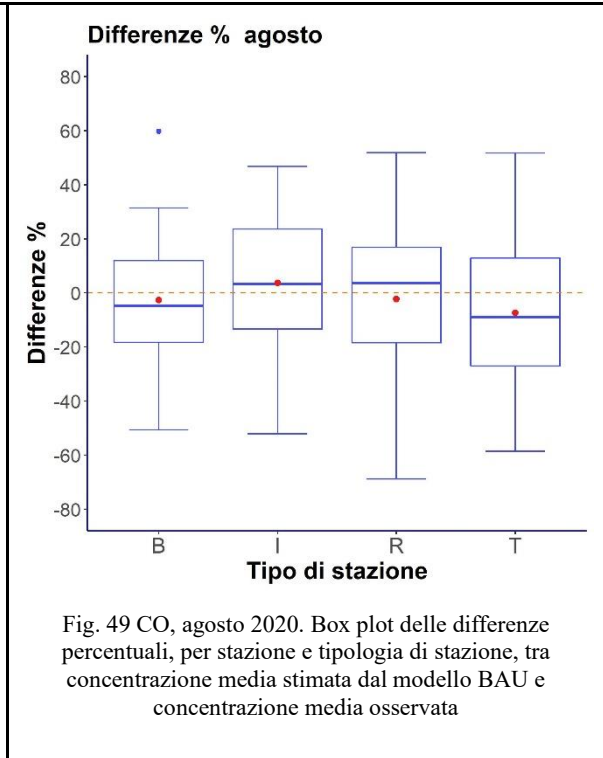
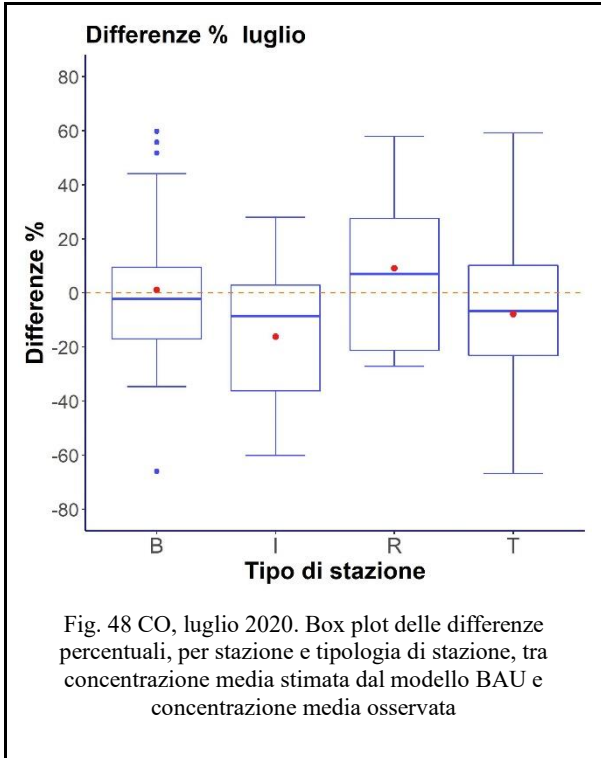


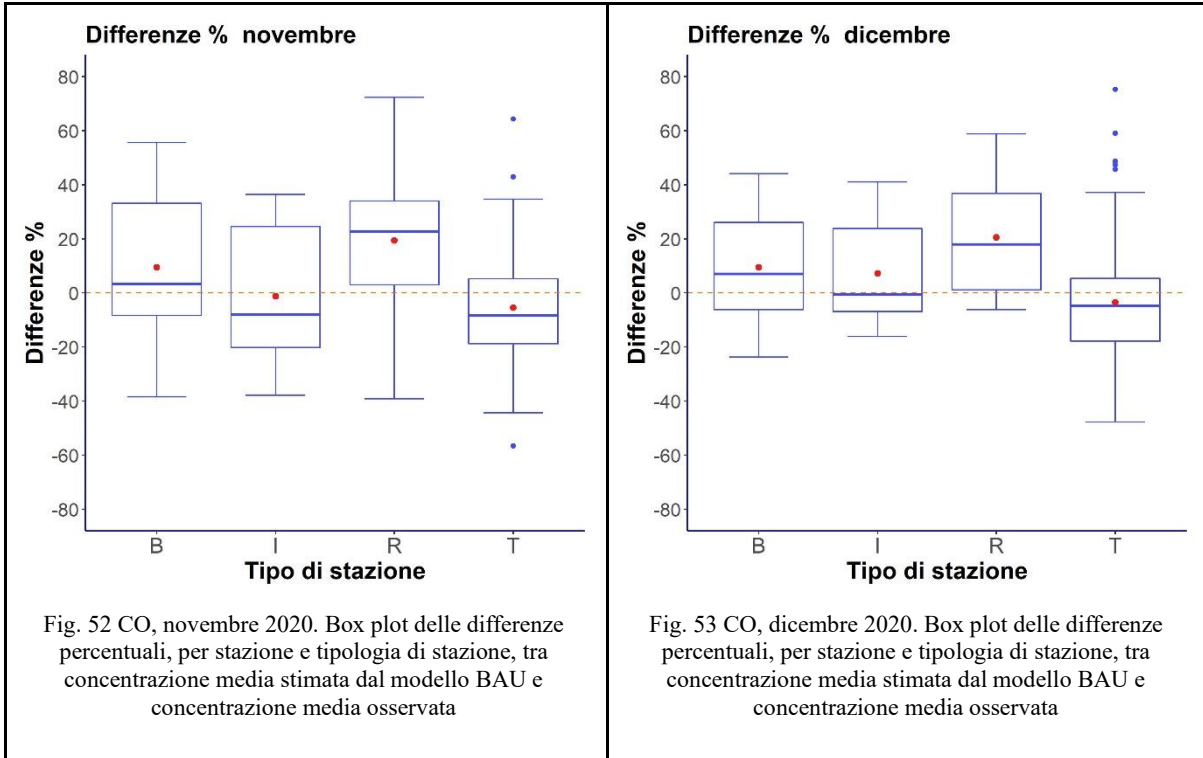
Fig. 43 Monossido di Carbonio Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)

Le differenze medie annue risultano negative per la quasi totalità delle stazioni, con valori osservati inferiori ai valori previsti dal modello, coerentemente con il fatto che durante il periodo in studio sono state messe in campo restrizioni agli spostamenti a causa della pandemia di coronavirus e con il fatto che il modello tende ad effettuare previsioni che mediamente sottostimano il dato osservato, come discusso nel paragrafo 3.2. Non si evidenzia un gradiente Nord-Sud che faccia ipotizzare un diverso comportamento delle stazioni in base alla latitudine.

I boxplot seguenti rappresentano la distribuzione, mese per mese, delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni, raggruppate per tipologia di stazione.







Per le stazioni rurali non si osserva una diminuzione delle concentrazioni durante i mesi analizzati, nemmeno durante i periodi di lockdown più restrittivo.

Le differenze osservate per le stazioni di background sono paragonabili a quelle osservate durante il periodo di validazione, oppure presentano variazioni nel senso opposto a quelle che ci si aspetterebbe, non evidenziando alcun effetto delle misure restrittive.

Le stazioni industriali presentano invece un calo delle differenze percentuali durante i mesi da marzo a luglio e nel mese di settembre, indicando una diminuzione della concentrazione di monossido di carbonio rispetto al valore atteso.

Un calo del valore osservato rispetto al valore previsto è osservato anche nelle stazioni da traffico per tutti i mesi dell'anno, con differenze più importanti (media intorno a -20%) nei mesi del lockdown primaverile (da marzo a maggio).

Tab. 7 Statistiche descrittive delle variazioni assolute di CO, nelle centraline valide, per tipo di stazione e per mese ( $\mu\text{g}/\text{m}^3$ )

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err
Marzo	-0,46	-0,10	-0,17	0,15	0,10	-0,79	-0,21	-0,16	0,31	0,24	-0,76	-0,45	-0,38	-0,46	-0,10
Aprile	-0,37	-0,15	-0,16	0,04	0,08	-0,84	-0,36	-0,45	0,19	0,22	-0,80	-0,45	-0,32	-0,37	-0,15
Maggio	-0,36	-0,19	-0,18	-0,05	0,07	-0,72	-0,23	-0,34	-0,06	0,22	-0,62	-0,41	-0,30	-0,36	-0,19
Giugno	-0,31	-0,17	-0,15	-0,02	0,07	-0,77	-0,02	-0,47	-0,05	0,22	-0,48	-0,30	-0,22	-0,31	-0,17
Luglio	-0,27	-0,12	-0,13	0,01	0,07	-0,90	-0,60	-0,72	-0,13	0,22	-0,40	-0,20	-0,16	-0,27	-0,12
Agosto	-0,29	-0,13	-0,14	0,00	0,07	-0,81	-0,46	-0,43	-0,11	0,23	-0,49	-0,22	-0,18	-0,29	-0,13
Settembre	-0,29	-0,11	-0,12	0,04	0,07	-0,42	-0,27	-0,21	-0,03	0,23	-0,43	-0,24	-0,18	-0,29	-0,11
Ottobre	-0,33	-0,10	-0,11	0,10	0,09	-0,31	-0,15	-0,05	0,07	0,18	-0,54	-0,25	-0,20	-0,33	-0,10
Novembre	-0,58	-0,22	-0,20	0,20	0,12	-0,31	-0,01	0,01	0,27	0,22	-0,94	-0,38	-0,26	-0,58	-0,22
Dicembre	-0,63	-0,21	-0,21	0,25	0,13	-0,22	0,00	0,04	0,42	0,24	-0,69	-0,12	-0,15	-0,63	-0,21

Tab. 8 Statistiche descrittive delle variazioni di CO, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-60,24	-13,17	-24,28	23,45	-72,26	0,65	-21,16	67,82	-82,64	-41,47	-48,14	-10,18
Aprile	-71,56	-25,57	-34,61	14,74	-88,20	-14,34	-38,97	48,12	-89,84	-44,11	-55,40	-12,48
Maggio	-86,20	-43,61	-50,80	-10,27	-92,31	-26,73	-41,19	17,80	-96,76	-56,26	-68,05	-33,72
Giugno	-78,37	-37,40	-45,50	-0,62	-107,95	-27,96	-56,20	13,74	-84,54	-44,01	-53,07	-15,13
Luglio	-74,10	-22,20	-37,06	7,55	-104,58	-49,20	-71,07	-12,02	-80,11	-31,95	-40,89	-2,62
Agosto	-72,15	-25,26	-39,06	9,30	-88,99	-31,18	-38,76	18,14	-82,68	-31,00	-42,99	1,09
Settembre	-66,04	-18,70	-29,02	16,72	-63,56	-21,44	-34,50	0,33	-73,48	-28,45	-32,54	8,22
Ottobre	-53,00	-7,56	-15,59	26,99	-57,72	-4,63	-15,18	25,75	-65,50	-17,13	-25,53	18,32
Novembre	-55,13	-4,83	-11,65	37,65	-50,93	18,47	3,16	67,06	-65,41	-18,38	-24,71	20,69
Dicembre	-50,25	-1,50	-10,56	37,14	-36,87	26,24	14,93	74,49	-58,50	-6,37	-14,83	31,57

Tab. 9 Significatività delle variazioni di CO<sub>2</sub> nelle centraline valide, per tipo di variazione e per mese (%)

mese	Δ non significativa			Δ significativa		
	Δ negativa	Δ positiva	Totale non significativa	Δ negativa	Δ positiva	Totale significativa
<b>Marzo</b>	12,68%	9,41%	22,09%	52,79%	25,13%	77,91%
<b>Aprile</b>	12,92%	7,67%	20,58%	50,99%	28,43%	79,42%
<b>Maggio</b>	12,72%	8,65%	21,37%	56,75%	21,88%	78,63%
<b>Giugno</b>	13,93%	9,50%	23,44%	50,38%	26,18%	76,56%
<b>Luglio</b>	15,39%	11,48%	26,87%	42,32%	30,81%	73,13%
<b>Agosto</b>	12,80%	11,75%	24,56%	41,53%	33,92%	75,44%
<b>Settembre</b>	14,66%	12,12%	26,78%	40,61%	32,61%	73,22%
<b>Ottobre</b>	13,02%	12,97%	25,99%	36,54%	37,47%	74,01%
<b>Novembre</b>	11,60%	11,45%	23,05%	42,22%	34,74%	76,95%
<b>Dicembre</b>	11,29%	11,20%	22,49%	36,85%	40,66%	77,51%
<b>Totale</b>	<b>13,10%</b>	<b>10,61%</b>	<b>23,71%</b>	<b>45,15%</b>	<b>31,14%</b>	<b>76,29%</b>

### 3.3.3 Biossido di azoto

Nelle mappe seguenti sono riportate le differenze medie percentuali tra il dato osservato per le singole stazioni su cui è stato rilevato il biossido di azoto e il dato previsto dal modello. Le tre mappe si riferiscono ai periodi 1 marzo – 30 aprile, 1 maggio – 30 settembre e 1 ottobre - 31 dicembre 2020. I dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.



Leaflet | © OpenStreetMap contributors, CC-BY-SA, Tiles © Esri — Source: US National Park Service, Map tiles by Sta Design, CC BY 3.0 — Map data © OpenStreetMap contributors

Fig. 54 Biossido di azoto Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)



Leaflet | © OpenStreetMap contributors, CC-BY-SA, Tiles © Esri — Source: US National Park Service, Map tiles by Sta Design, CC BY 3.0 — Map data © OpenStreetMap contributors

Fig. 55 Biossido di azoto Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)

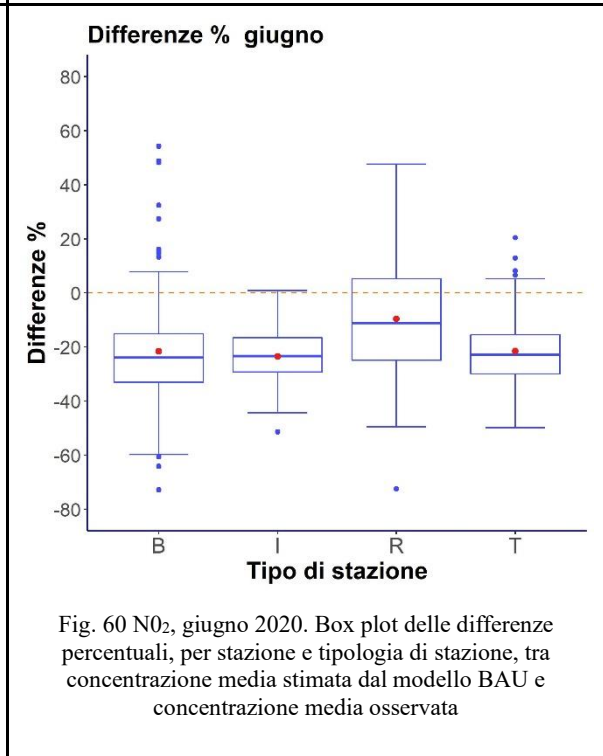
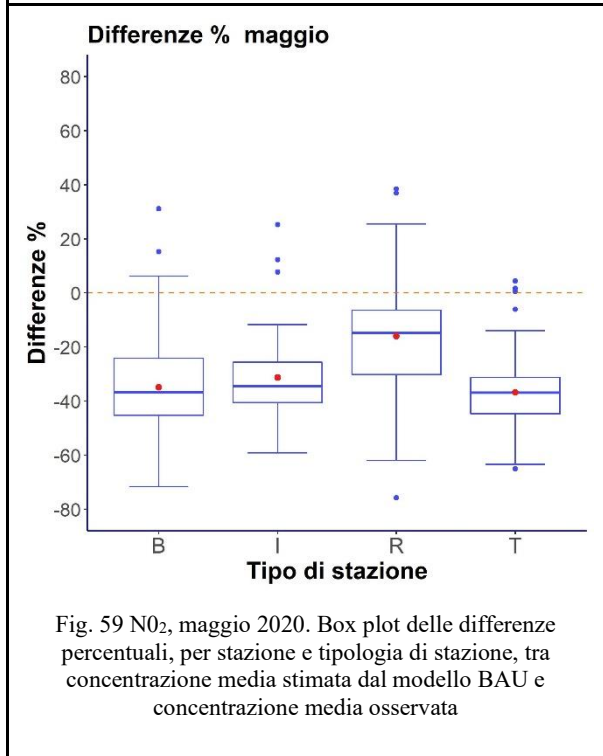
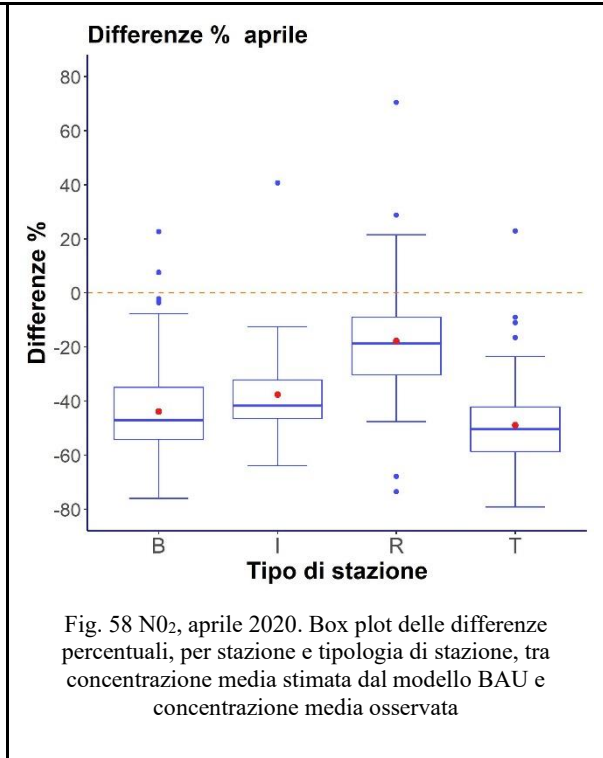
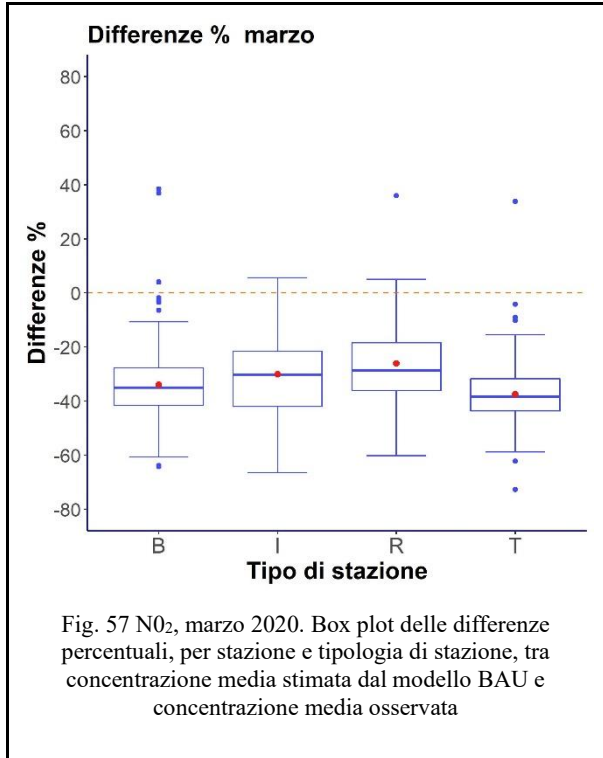


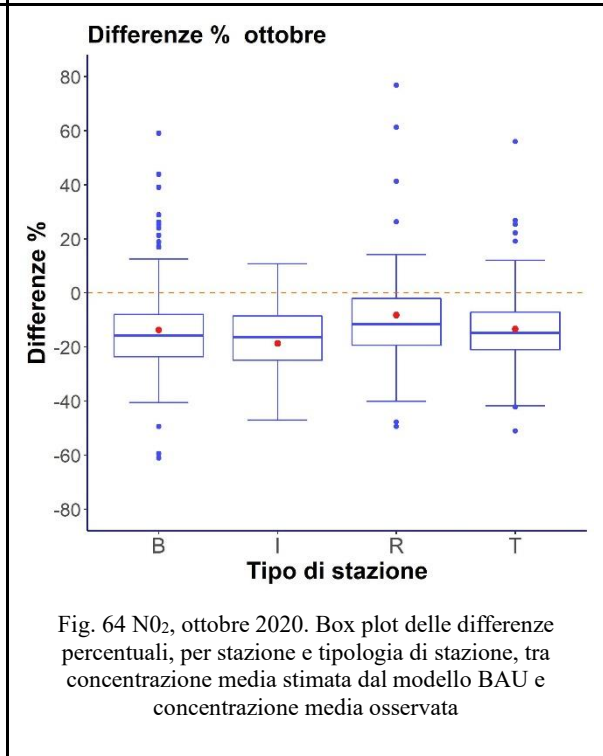
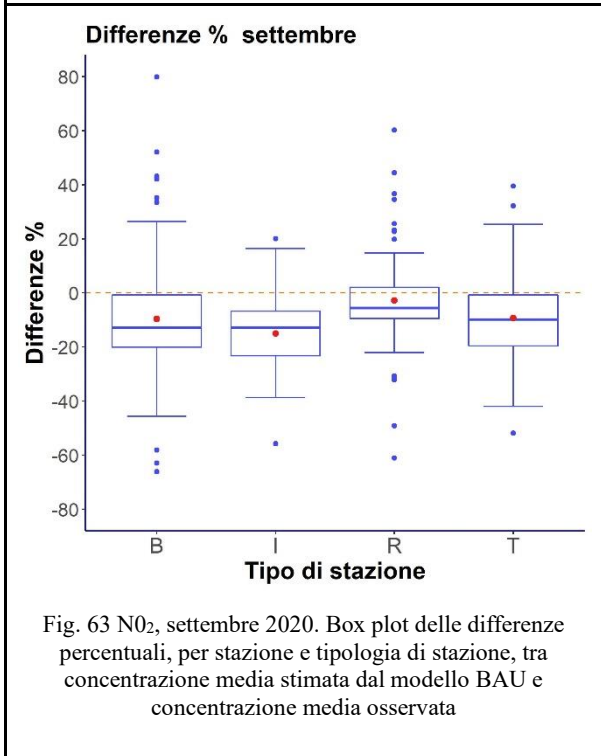
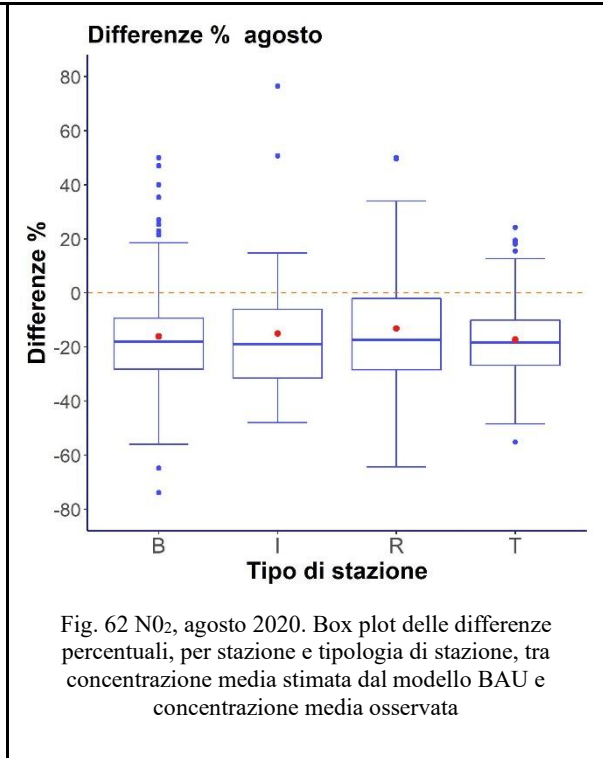
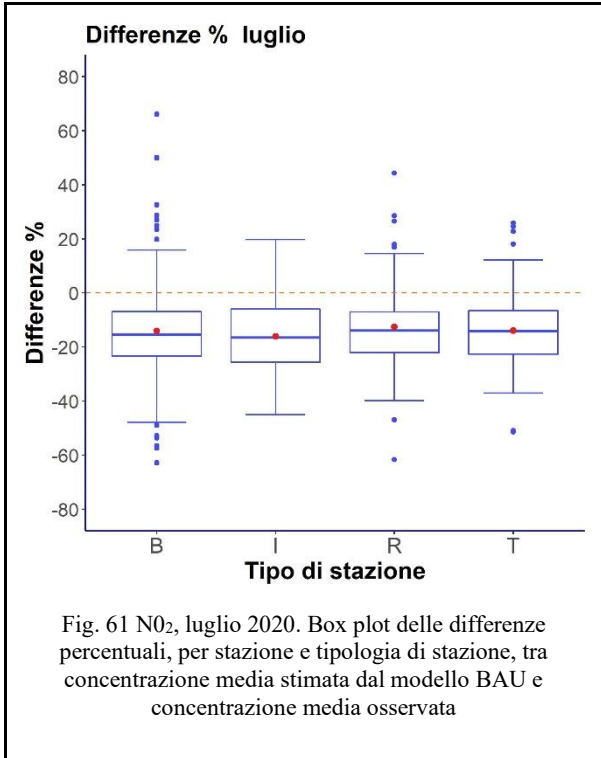


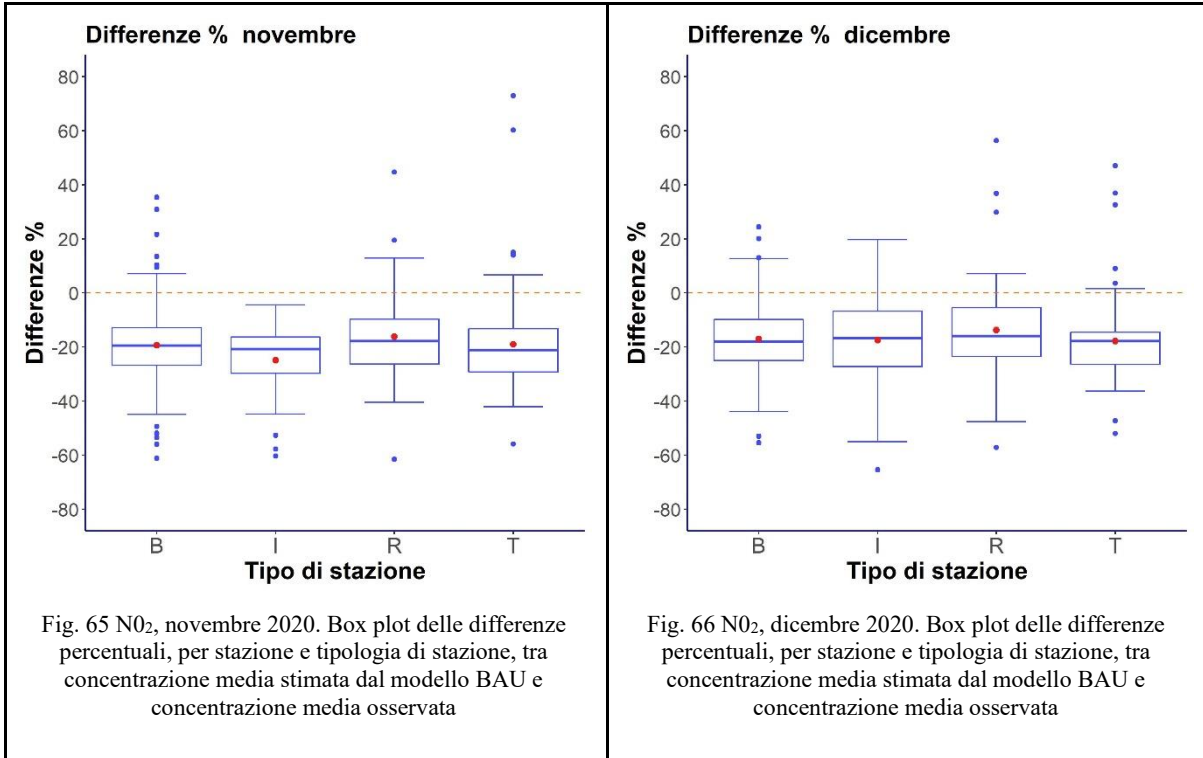
Fig. 56 Biossido di azoto Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)

Tutte le stazioni presentano differenze medie annue negative, con valori osservati inferiori ai valori previsti dal modello, coerentemente con il fatto che durante il periodo in studio sono state messe in campo restrizioni agli spostamenti a causa della pandemia di coronavirus e con il fatto che il modello tende ad effettuare previsioni che mediamente sottostimano il dato osservato, come discusso nel paragrafo 3.2. Non si evidenzia un gradiente Nord-Sud che faccia ipotizzare un diverso comportamento delle stazioni in base alla latitudine.

I boxplot seguenti rappresentano la distribuzione, mese per mese, delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni, raggruppate per tipologia di stazione.







Il biossido di azoto evidenzia un chiaro calo durante i mesi del lockdown primaverile e nel resto dell'anno.

Durante il mese di marzo e maggio, media e mediana delle differenze si attestano tra il -30% e -40% circa per le 4 tipologie di stazioni, con un calo leggermente inferiore nelle stazioni rurali. Un calo ancora maggiore si evidenzia durante il mese di aprile, con la media delle differenze inferiore a -40% per le stazioni di background, industriali e da traffico, mentre per le sole stazioni di tipo rurale la media delle differenze è intorno al -20%.

Durante i mesi da giugno a dicembre, la media delle differenze si attesta, in maniera abbastanza uniforme tra le diverse tipologie di stazioni, tra -10% e -20% (durante il periodo di validazione le differenze medie non superavano il -10% per nessuna tipologia di stazione).

Tab. 10 Statistiche descrittive delle variazioni di NO<sub>2</sub>, nelle centraline valide, per tipo di stazione e per mese (µg/m<sup>3</sup>)

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err
Marzo	-11,07	-7,11	-6,35	-2,33	1,20	-7,94	-4,83	-3,74	-1,05	1,07	-19,32	-13,16	-12,64	-6,42	1,59
Aprile	-10,53	-7,29	-6,46	-2,79	1,08	-8,35	-5,44	-4,73	-1,96	1,00	-20,93	-15,64	-14,50	-9,24	1,52
Maggio	-7,11	-4,62	-3,87	-1,37	1,00	-6,14	-3,86	-3,55	-0,97	0,94	-14,09	-9,90	-9,10	-5,00	1,43
Giugno	-4,73	-2,73	-2,25	-0,34	0,97	-5,15	-2,92	-2,56	-0,45	0,93	-8,88	-5,32	-4,81	-1,46	1,38
Luglio	-3,91	-1,97	-1,70	0,07	0,98	-4,37	-2,10	-1,88	0,27	0,93	-6,82	-3,50	-3,16	-0,05	1,39
Agosto	-4,16	-2,21	-1,95	-0,13	0,99	-4,35	-1,69	-1,77	0,62	0,93	-8,15	-4,28	-4,06	-0,52	1,43
Settembre	-3,99	-1,61	-1,38	0,81	1,02	-4,03	-1,45	-1,48	0,75	0,92	-6,88	-2,74	-2,71	1,40	1,45
Ottobre	-5,61	-2,58	-2,42	0,31	1,12	-5,09	-1,96	-2,10	0,51	0,98	-8,66	-4,11	-4,15	0,11	1,51
Novembre	-8,75	-5,05	-4,40	-0,91	1,26	-7,59	-4,66	-3,87	-0,95	1,01	-13,41	-7,70	-7,58	-2,50	1,64
Dicembre	-8,90	-4,65	-3,94	-0,12	1,33	-7,27	-3,60	-2,88	0,42	1,16	-12,69	-6,86	-6,51	-1,00	1,64

Tab. 11 Statistiche descrittive delle variazioni di NO<sub>2</sub>, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-51,87	-32,16	-35,26	-16,30	-48,82	-26,77	-29,68	-8,36	-57,10	-37,49	-39,66	-20,79
Aprile	-58,11	-39,02	-42,95	-24,66	-53,79	-32,75	-35,65	-19,14	-64,31	-49,67	-51,40	-37,59
Maggio	-51,42	-31,50	-33,92	-15,20	-48,06	-27,08	-30,61	-11,26	-52,43	-37,20	-37,39	-22,46
Giugno	-38,51	-19,52	-22,29	-3,95	-39,17	-20,18	-24,08	-5,50	-37,42	-21,40	-22,71	-7,35
Luglio	-32,26	-13,37	-16,08	0,64	-34,68	-13,83	-17,21	2,80	-29,62	-14,21	-14,76	-0,43
Agosto	-35,02	-15,85	-18,99	-1,54	-35,49	-8,43	-14,48	6,91	-34,31	-16,97	-18,58	-2,60
Settembre	-28,16	-9,07	-11,19	6,99	-30,62	-6,56	-12,50	7,78	-26,42	-9,32	-10,83	5,39
Ottobre	-32,02	-13,02	-14,97	1,95	-36,79	-9,08	-15,98	5,22	-28,19	-11,67	-13,77	0,30
Novembre	-34,75	-18,92	-19,72	-4,95	-43,34	-22,31	-23,18	-6,68	-35,31	-19,07	-20,63	-7,36
Dicembre	-33,51	-16,34	-17,30	-0,67	-39,78	-17,06	-17,73	2,83	-32,60	-16,61	-16,91	-2,68

Tab. 12 Significatività delle variazioni di NO<sub>2</sub>, nelle centraline valide, per tipo di variazione e per mese (%)

mese	Δ non significativa			Δ significativa		Totale significativa
	Δ negativa	Δ positiva	Totale non significativa	Δ negativa	Δ positiva	
<b>Marzo</b>	9,06%	4,70%	13,76%	80,22%	6,02%	86,24%
<b>Aprile</b>	6,13%	3,16%	9,29%	87,29%	3,42%	90,71%
<b>Maggio</b>	11,37%	5,77%	17,14%	77,71%	5,15%	82,86%
<b>Giugno</b>	17,35%	10,49%	27,84%	62,85%	9,31%	72,16%
<b>Luglio</b>	22,01%	12,69%	34,70%	52,37%	12,92%	65,30%
<b>Agosto</b>	19,72%	11,42%	31,14%	56,30%	12,56%	68,86%
<b>Settembre</b>	18,99%	14,03%	33,02%	47,97%	19,01%	66,98%
<b>Ottobre</b>	16,29%	10,65%	26,94%	56,65%	16,42%	73,06%
<b>Novembre</b>	11,83%	7,57%	19,40%	69,60%	11,00%	80,60%
<b>Dicembre</b>	13,60%	9,77%	23,37%	62,65%	13,98%	76,63%
<b>Totale</b>	<b>14,59%</b>	<b>8,98%</b>	<b>23,57%</b>	<b>65,54%</b>	<b>10,89%</b>	<b>76,43%</b>

### 3.3.4 Ozono

Vengono di seguito riportati i grafici relativi alla massima media oraria, che è il parametro che presentava la media delle differenze percentuali più prossima allo zero durante il periodo di validazione. I risultati per la massima media mobile di 8 h sono comunque disponibili in appendice e risultano molto simili a quelli presentati di seguito riferiti alla massima media oraria.

Nelle cartine seguenti sono riportate le differenze medie percentuali tra il dato previsto dal modello e il dato osservato per le singole stazioni su cui è stato rilevato l'ozono. Le tre mappe si riferiscono ai periodi 1 marzo – 30 aprile, 1 maggio – 30 settembre e 1 ottobre - 31 dicembre 2020.

I dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.

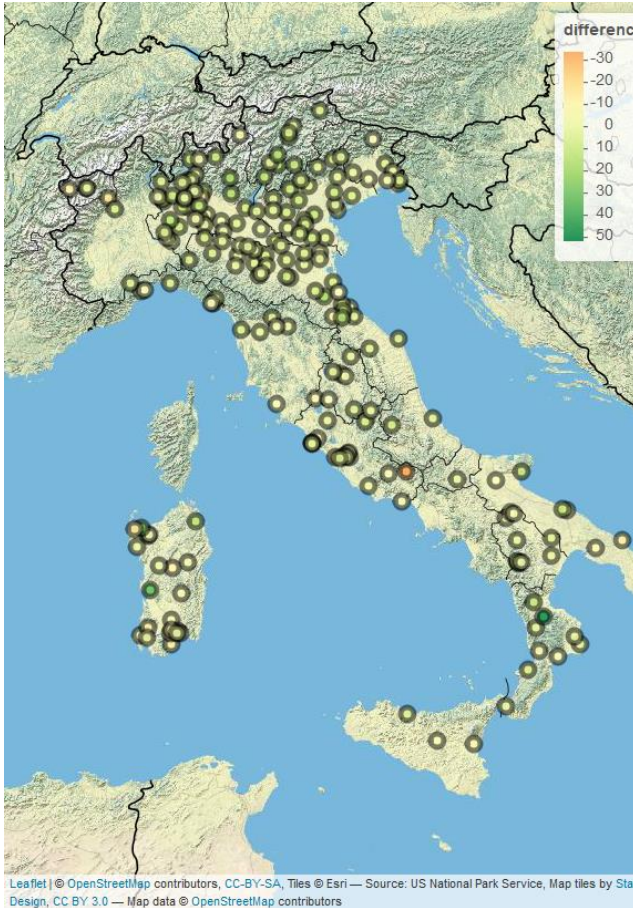


Fig. 67 Ozono: massima media oraria Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)

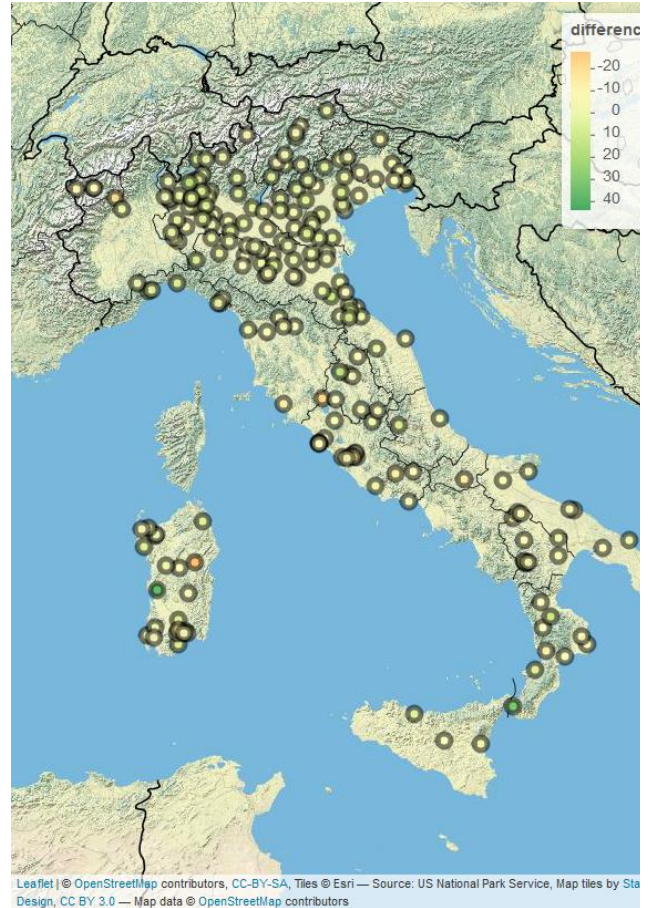


Fig. 68 Ozono: massima media oraria Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)

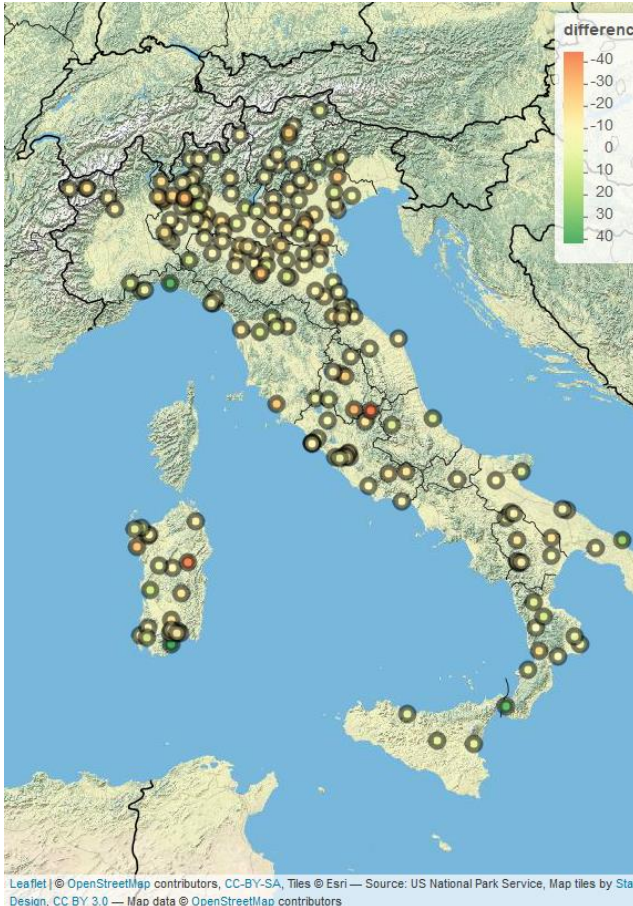
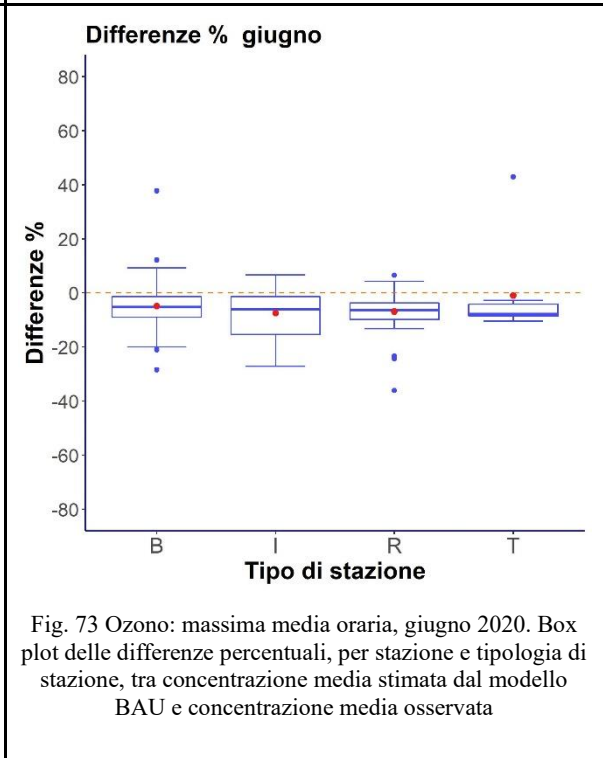
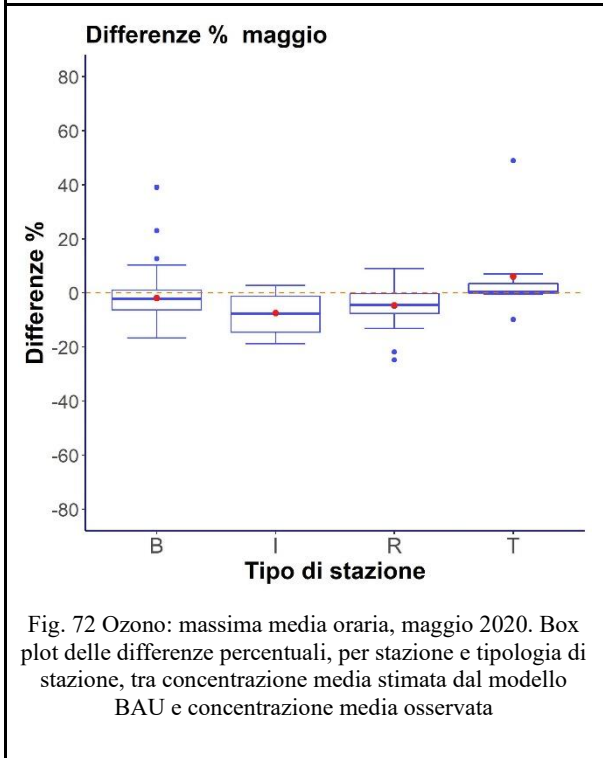
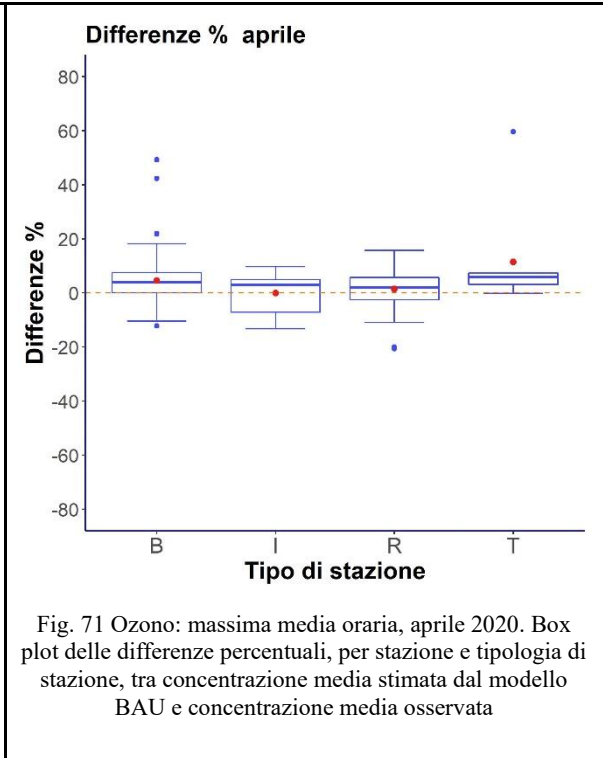
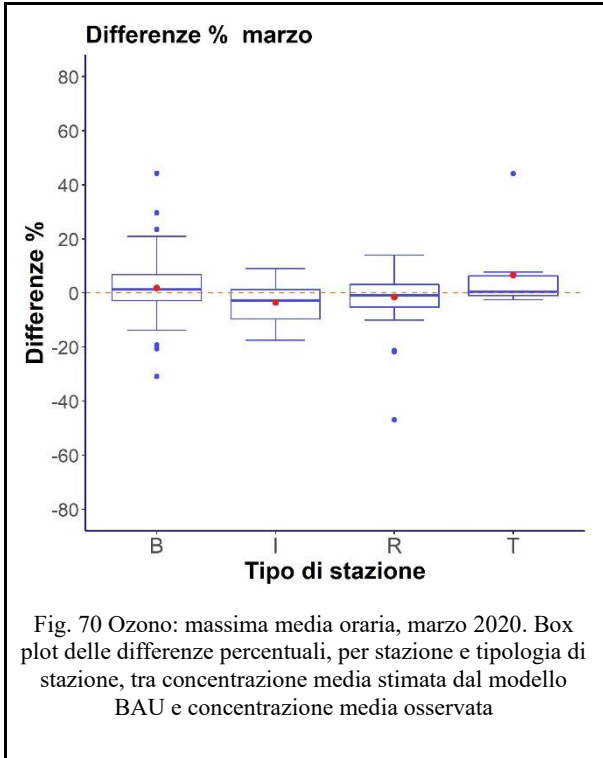


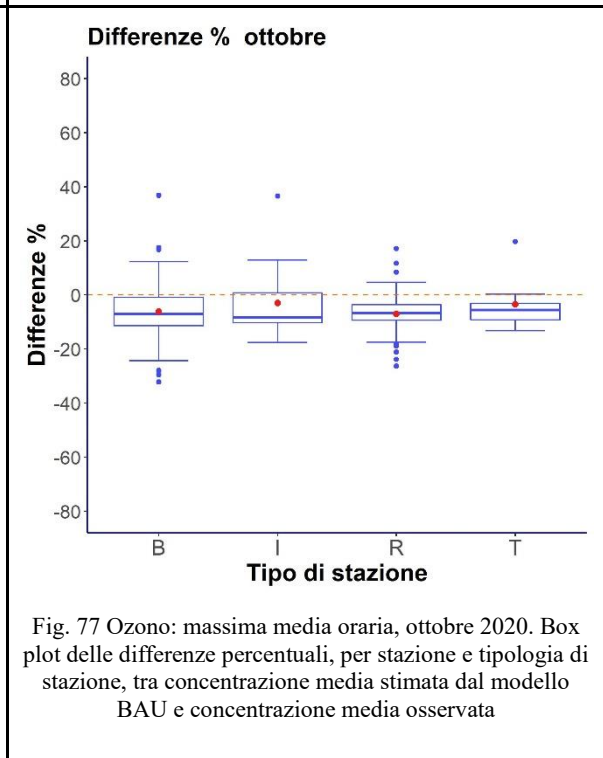
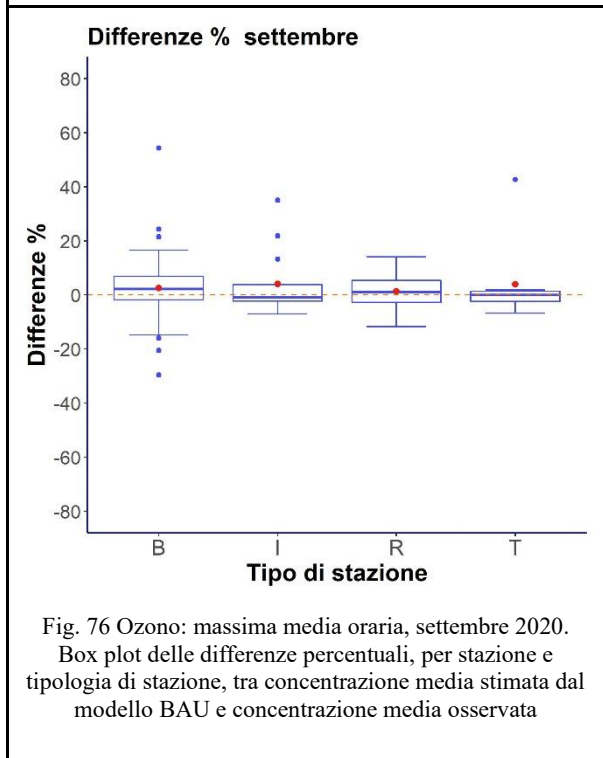
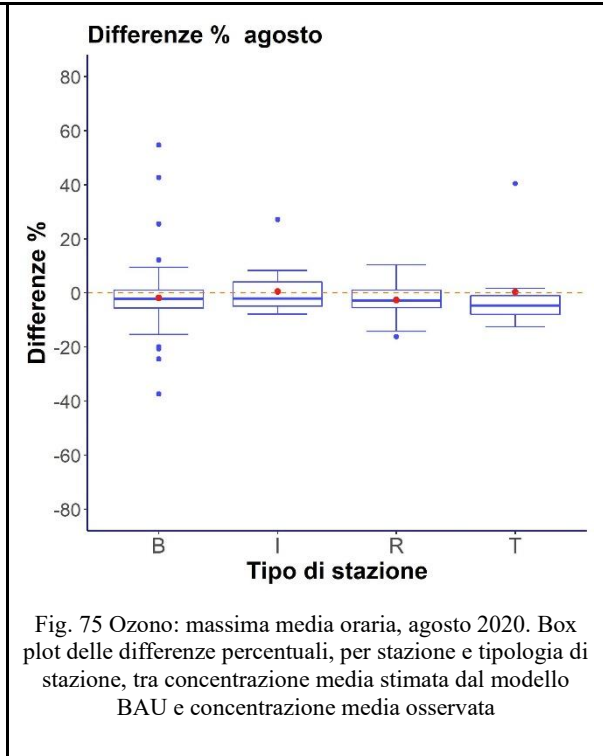
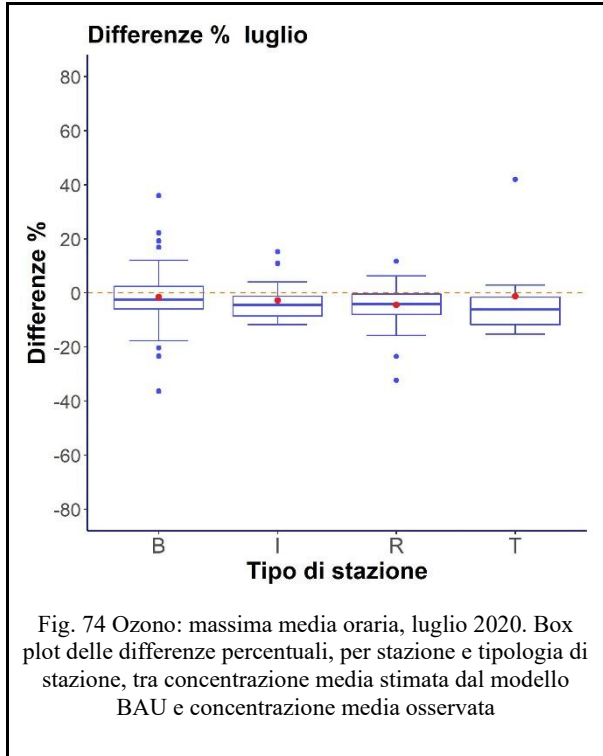
Fig. 69 Ozono: massima media oraria Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)

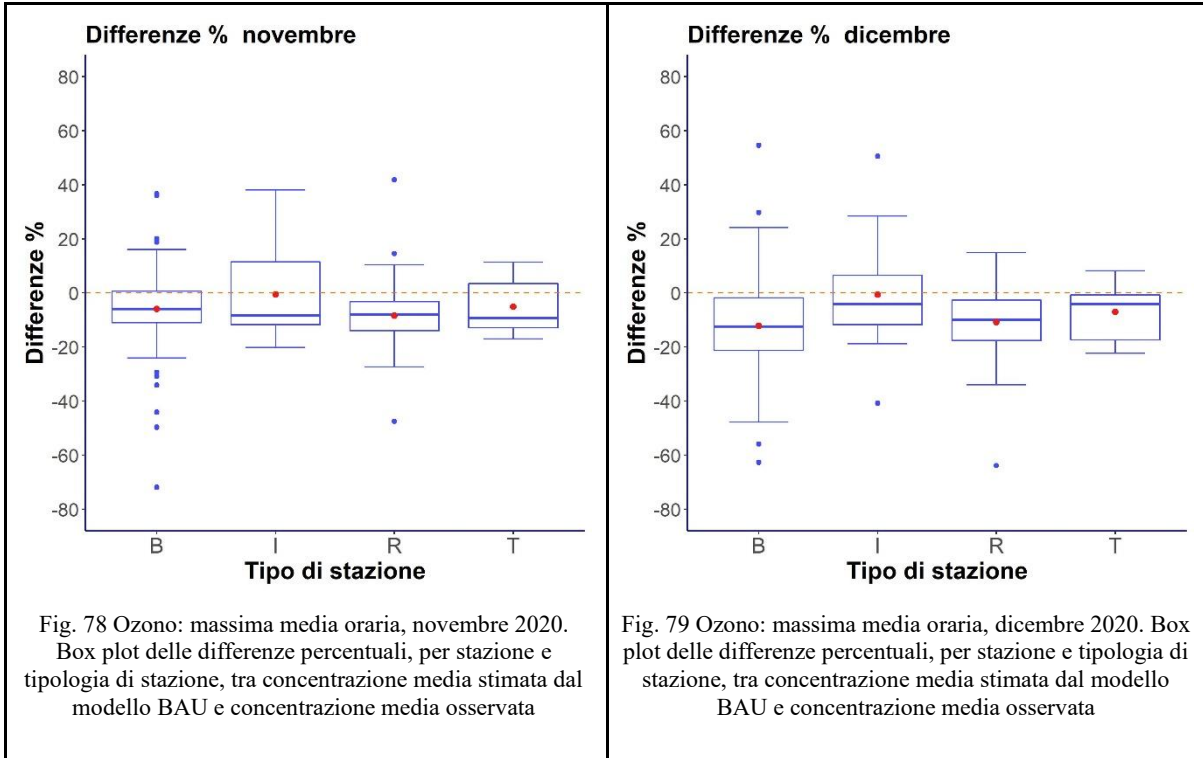
Nella maggior parte delle stazioni le differenze medie tra valori osservati e valori previsti dal modello sono di piccola entità e non sembra evidenziarsi una tendenza né all'aumento né alla diminuzione in tutti e tre i periodi considerati. Anche in questo caso non si evidenzia un gradiente Nord-Sud che faccia ipotizzare un diverso comportamento delle stazioni in base alla latitudine.

I boxplot seguenti rappresentano la distribuzione, mese per mese, delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni, raggruppate per tipologia di stazione.









Per quel che riguarda la massima media oraria dell'ozono, le differenze percentuali nei mesi da marzo a dicembre 2020 sono paragonabili a quelle osservate nel periodo di validazione, non evidenziando alcuna variazione rispetto agli anni precedenti, per nessuna tipologia di stazione.

Tab. 13 Statistiche descrittive delle variazioni di Ozono, nelle centraline valide, per tipo di stazione e per mese ( $\mu\text{g}/\text{m}^3$ )

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err
Marzo	-7,71	-0,30	-0,76	6,92	2,78	-9,30	-1,75	-2,05	4,33	2,60	-3,88	3,61	1,43	8,54	2,51
Aprile	-5,62	3,46	2,71	11,92	2,79	-7,82	1,55	0,74	9,45	2,57	-4,45	7,74	5,45	15,73	2,57
Maggio	-11,55	-2,88	-3,39	5,66	2,83	-13,83	-5,89	-6,02	2,63	2,57	-6,58	2,94	0,73	10,47	2,54
Giugno	-15,23	-6,20	-6,14	2,90	2,99	-14,93	-6,72	-7,08	1,69	2,53	-12,02	-3,44	-5,45	2,96	2,65
Luglio	-13,72	-2,70	-3,97	6,92	3,13	-13,25	-3,32	-5,02	5,49	2,76	-14,10	-2,77	-5,02	5,47	3,00
Agosto	-12,32	-2,50	-3,52	6,25	3,14	-10,27	-1,53	-2,23	7,39	2,64	-13,00	-0,87	-3,11	7,40	2,67
Settembre	-7,56	2,38	1,37	11,10	3,00	-5,87	2,91	2,04	10,34	2,54	-7,45	2,34	1,38	9,93	2,62
Ottobre	-12,72	-4,24	-4,75	3,85	2,83	-12,22	-3,79	-5,34	3,16	2,35	-11,28	-2,63	-3,43	5,20	2,46
Novembre	-10,59	-3,03	-2,99	4,75	2,79	-13,12	-4,07	-5,51	3,59	2,43	-10,34	-3,16	-3,91	4,33	2,49
Dicembre	-10,69	-3,76	-4,05	3,88	2,77	-9,44	-1,78	-2,42	3,95	2,31	-8,37	-2,32	-3,24	4,18	2,45

Tab. 14 Statistiche descrittive delle variazioni di Ozono, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-8,99	0,85	-0,90	8,54	-9,88	-1,50	-2,18	4,83	-4,89	6,68	1,64	11,25
Aprile	-5,60	3,62	2,62	11,43	-8,26	1,06	0,76	9,48	-4,63	11,21	5,67	16,34
Maggio	-11,09	-2,68	-3,23	5,34	-13,73	-5,96	-5,82	2,52	-6,68	6,13	0,70	11,23
Giugno	-13,96	-5,48	-5,71	2,63	-14,67	-6,61	-6,56	1,58	-12,68	-1,22	-5,41	2,89
Luglio	-11,67	-2,46	-3,39	5,70	-11,43	-3,02	-4,31	4,79	-13,17	-1,16	-5,01	5,04
Agosto	-10,90	-2,15	-3,02	5,31	-9,66	-0,73	-1,92	6,34	-13,83	0,34	-3,19	7,84
Settembre	-7,78	2,18	1,36	10,86	-5,95	3,55	2,00	10,41	-8,23	4,00	1,84	11,77
Ottobre	-18,04	-6,19	-6,65	5,54	-14,07	-3,93	-6,60	3,63	-15,88	-3,37	-5,24	6,62
Novembre	-22,73	-6,08	-5,61	10,11	-17,54	-4,69	-7,69	5,42	-19,05	-5,00	-6,83	8,07
Dicembre	-35,18	-12,00	-8,63	7,97	-13,63	-3,55	-3,51	5,99	-16,96	-7,04	-6,07	6,58

Tab. 15 Significatività delle variazioni di Ozono, nelle centraline valide, per tipo di variazione e per mese (%)

mese	$\Delta$ non significativa			$\Delta$ significativa		Totale significativa
	$\Delta$ negativa	$\Delta$ positiva	Totale non significativa	$\Delta$ negativa	$\Delta$ positiva	
Marzo	19,45%	18,54%	37,99%	33,66%	28,35%	62,01%
Aprile	14,85%	16,54%	31,40%	26,82%	41,79%	68,60%
Maggio	17,09%	14,13%	31,22%	43,78%	25,00%	68,78%
Giugno	16,39%	12,44%	28,83%	52,24%	18,93%	71,17%
Luglio	15,03%	12,59%	27,62%	45,74%	26,64%	72,38%
Agosto	17,32%	14,55%	31,87%	42,50%	25,63%	68,13%
Settembre	15,99%	15,52%	31,51%	30,12%	38,36%	68,49%
Ottobre	16,70%	13,42%	30,12%	48,30%	21,58%	69,88%
Novembre	18,31%	16,20%	34,50%	42,74%	22,76%	65,50%
Dicembre	18,72%	14,20%	32,92%	44,97%	22,11%	67,08%
Totale	16,99%	14,82%	31,80%	41,09%	27,11%	68,20%

### 3.3.5 Particolato atmosferico: PM<sub>10</sub>

Nelle cartine seguenti sono riportate le differenze medie percentuali tra il dato osservato e il dato previsto dal modello per le singole stazioni su cui è stato rilevato il PM<sub>10</sub>. Le tre mappe si riferiscono ai periodi 1 marzo – 30 aprile, 1 maggio – 30 settembre e 1 ottobre - 31 dicembre 2020.

I dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.

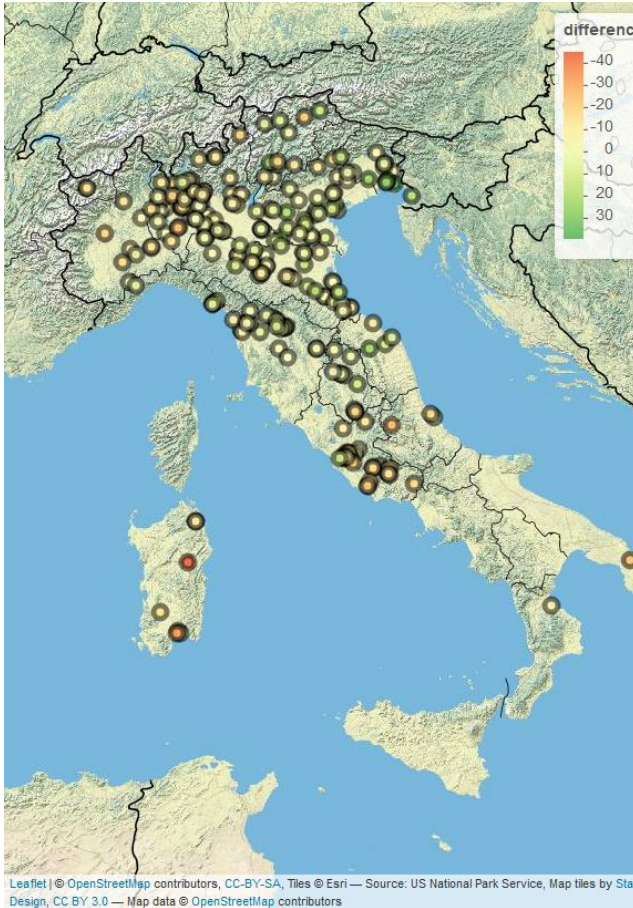


Fig. 80 PM<sub>10</sub> Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)

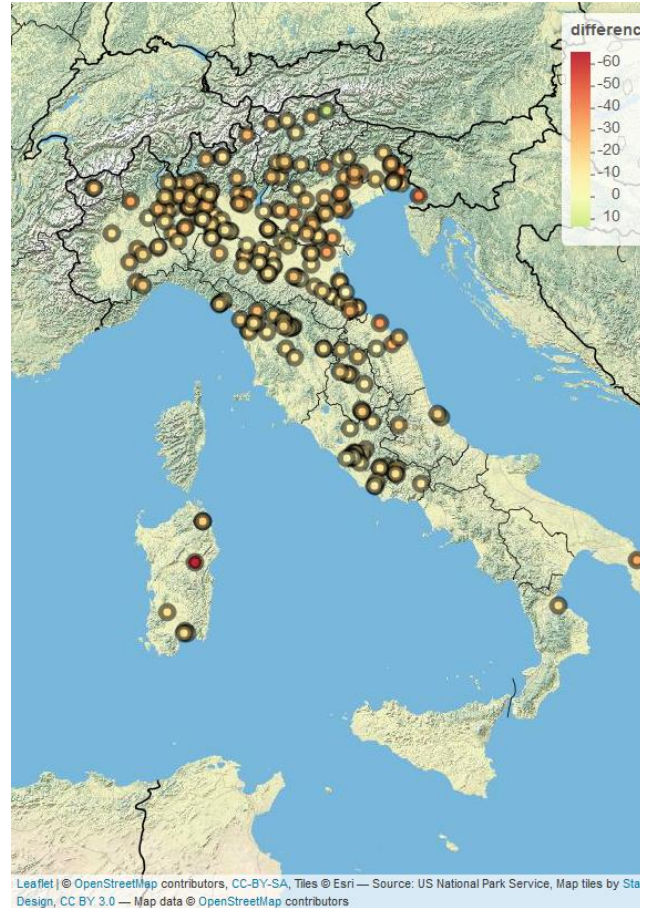


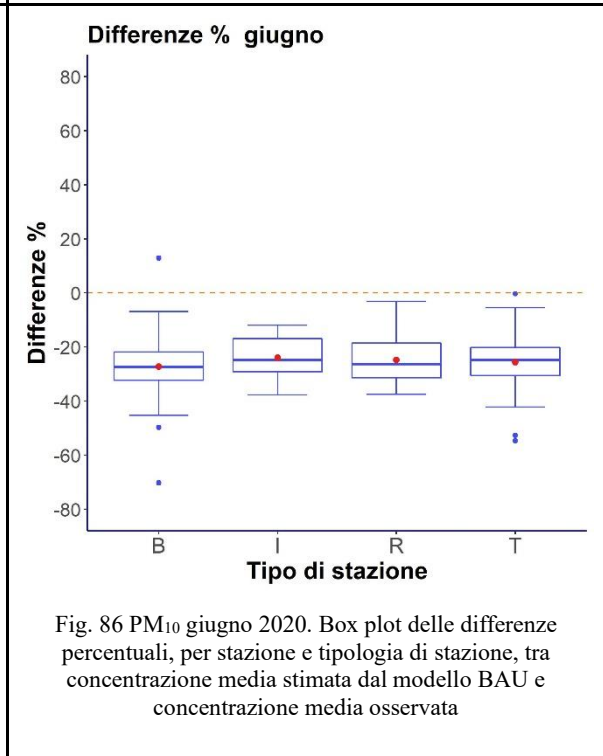
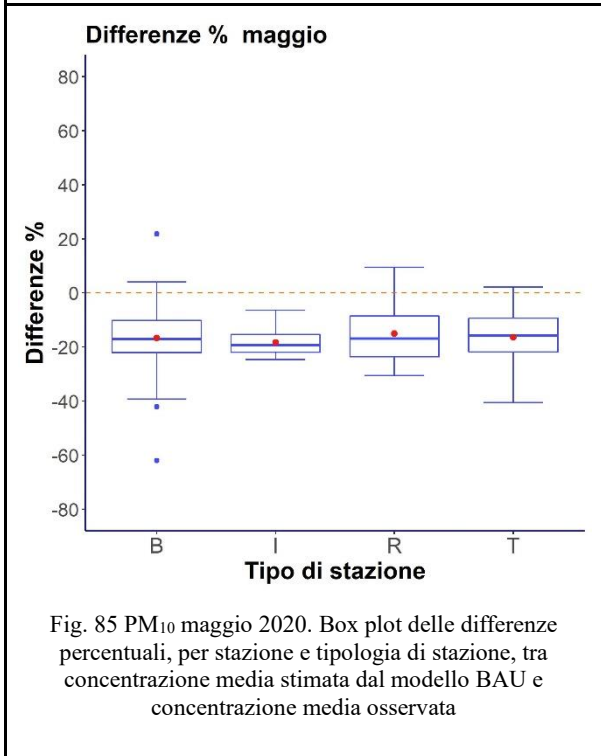
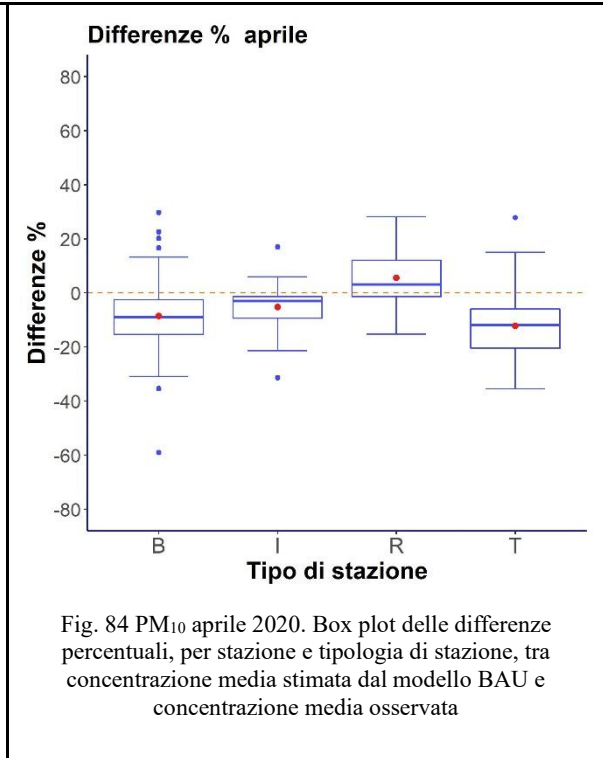
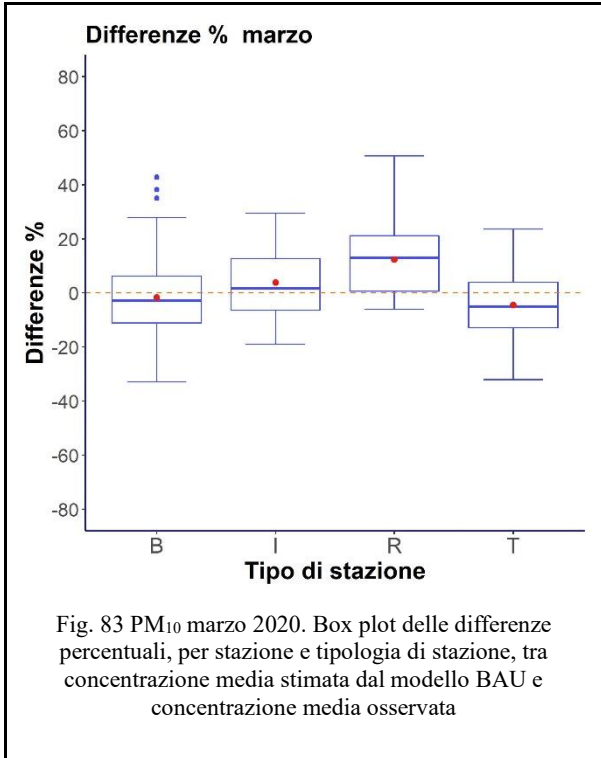
Fig. 81 PM<sub>10</sub> Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)



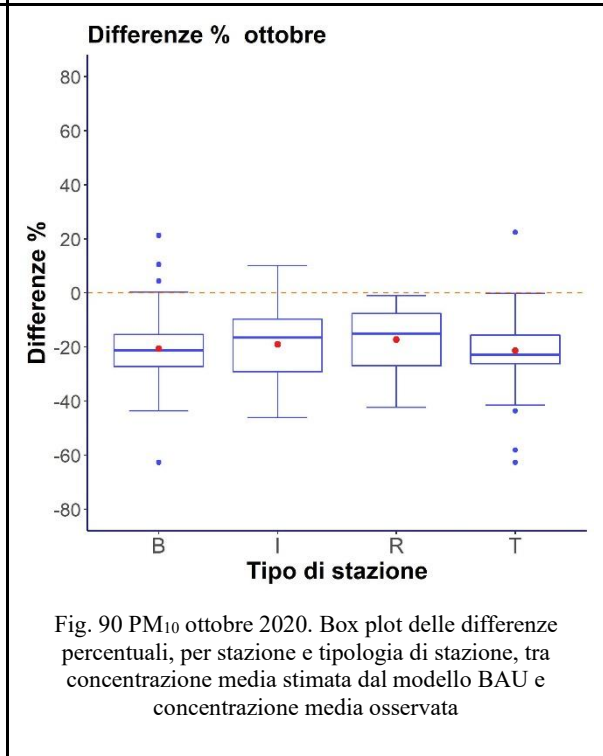
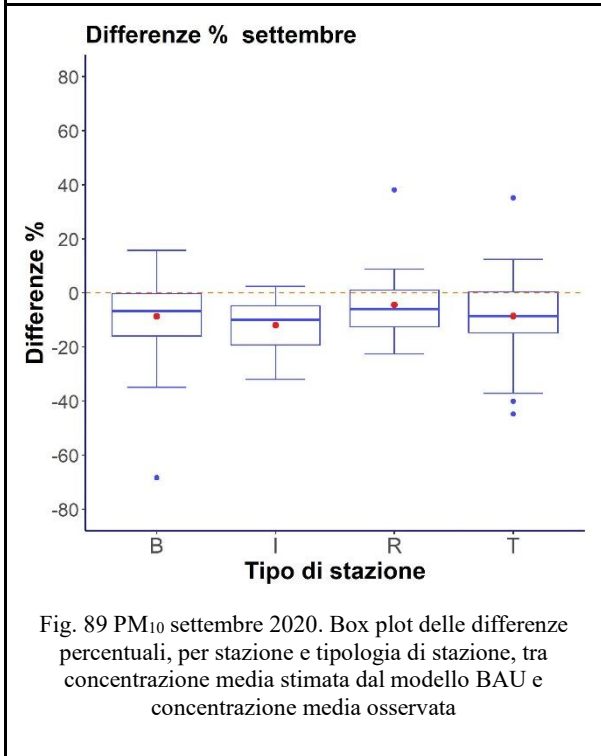
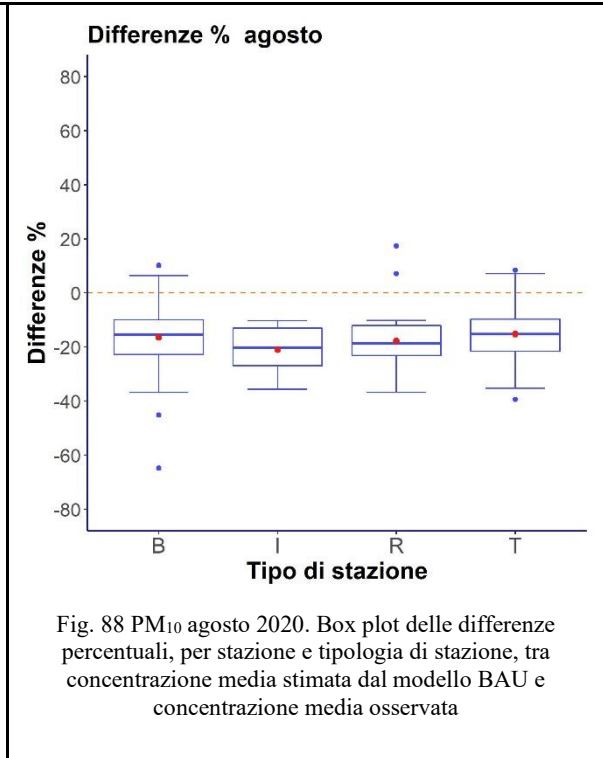
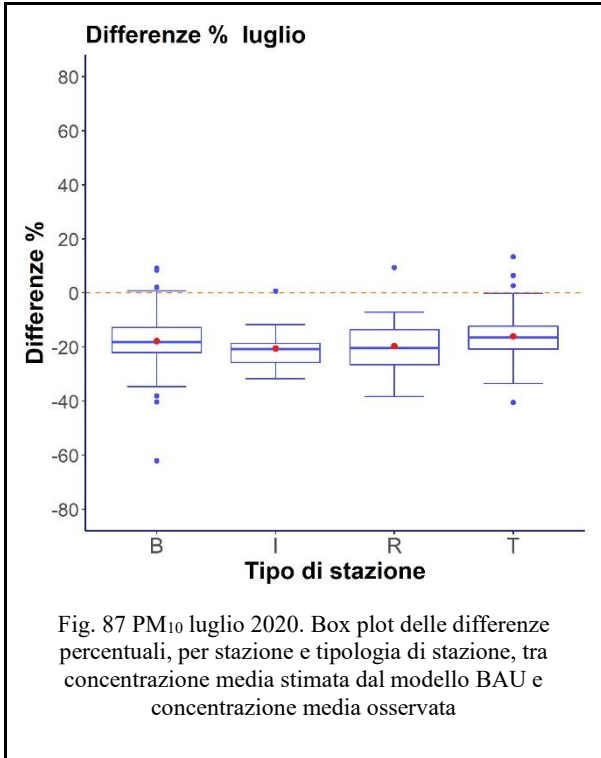
Fig. 82 PM<sub>10</sub> Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)

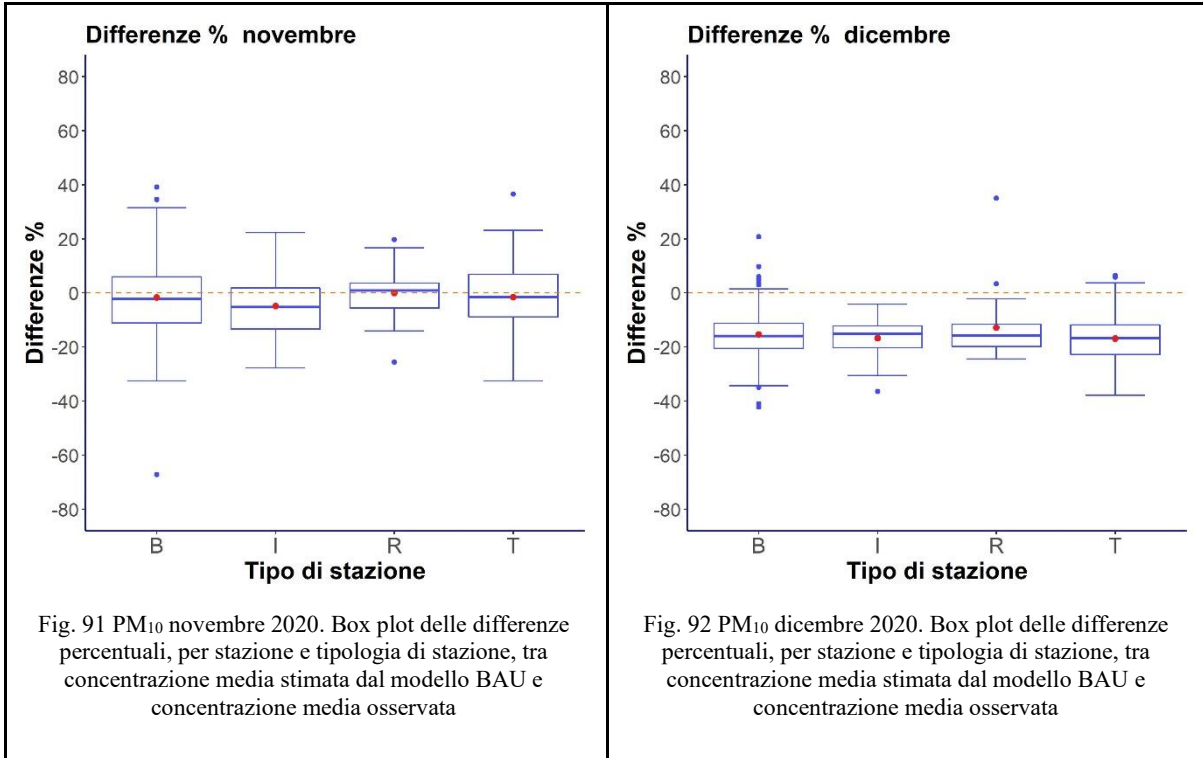
Le differenze medie annue risultano negative per la quasi totalità delle stazioni, con valori osservati inferiori ai valori previsti dal modello, coerentemente con il fatto che durante il periodo in studio sono state messe in campo restrizioni agli spostamenti a causa della pandemia di coronavirus e con il fatto che il modello tende ad effettuare previsioni che mediamente sottostimano il dato osservato, come discusso nel paragrafo 3.2. Non si evidenzia un gradiente Nord-Sud che faccia ipotizzare un diverso comportamento delle stazioni in base alla latitudine.

I boxplot seguenti rappresentano la distribuzione, mese per mese, delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni, raggruppate per tipologia di stazione.









Per i mesi di marzo e aprile non si rilevano differenze medie percentuali molto diverse da quelle riscontrate durante il periodo di validazione. Il mese di marzo è stato caratterizzato da fenomeni di trasporto di polvere a grande scala provenienti dall'Est Europa e che non possono essere previsti dal modello, che potrebbero aver compensato un eventuale calo che si fosse verificato a seguito del calo di emissioni antropiche. Nei mesi di maggio, giugno, luglio e agosto, ottobre e dicembre si osservano invece differenze medie percentuali intorno a -20%, a fronte di differenze medie entro -10% osservate durante il periodo di validazione. Le differenze sono abbastanza omogenee per le 4 tipologie di stazioni. Si può quindi affermare che, nonostante una certa variabilità da un mese all'altro, dovuta in parte a fenomeni di trasporto non prevedibili dal modello, il PM<sub>10</sub> ha mostrato una diminuzione nei mesi successivi a febbraio 2020, probabilmente imputabili alle misure restrittive imposte per il contenimento della pandemia.

Tab. 16 Statistiche descrittive delle variazioni di PM<sub>10</sub>, nelle centraline valide, per tipo di stazione e per mese (µg/m<sup>3</sup>)

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err.	25th	Media	Mediana	75th	Std. err.	25th	Media	Mediana	75th	Std. err.
Marzo	-9,85	-0,91	-4,55	1,06	2,17	10,59	1,35	-4,40	3,02	2,30	-11,33	-2,27	-5,48	0,29	2,31
Aprile	-5,94	-1,66	-1,79	2,53	1,85	-4,76	-0,20	-0,46	4,09	2,00	-7,83	-3,21	-3,20	1,04	2,00
Maggio	-7,08	-3,07	-3,85	0,04	1,74	-7,20	-3,05	-3,56	0,17	1,83	-7,83	-3,42	-4,54	0,18	1,84
Giugno	-7,36	-4,77	-4,73	2,12	1,70	-7,47	-4,84	-4,63	2,00	1,81	-7,85	-5,09	-5,13	2,38	1,82
Luglio	-6,33	-3,34	-3,67	0,76	1,68	-7,18	-4,31	-4,42	1,89	1,84	-6,45	-3,26	-3,69	0,49	1,79
Agosto	-6,03	-3,07	-3,13	0,04	1,71	-8,13	-4,82	-4,76	1,78	1,90	-6,45	-3,23	-3,46	0,14	1,83
Settembre	-6,06	-1,20	-2,13	3,48	1,76	-7,74	-2,65	-3,54	2,43	1,99	-6,75	-1,69	-2,62	3,70	1,88
Ottobre	-9,27	-4,57	-5,19	1,46	2,07	-11,80	-4,95	-7,21	0,85	2,40	10,27	-5,68	-5,98	1,73	2,19
Novembre	-8,34	0,21	-1,89	6,60	2,41	-11,19	-0,74	-3,40	7,82	2,78	-9,50	-0,04	-2,39	7,88	2,58
Dicembre	-10,12	-4,15	-4,81	1,06	2,48	12,37	-5,35	-6,61	0,92	2,85	-11,42	-4,64	-5,77	0,95	2,70

Tab. 17 Statistiche descrittive delle variazioni di PM<sub>10</sub>, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-37,20	-0,20	-18,98	4,47	-35,67	10,26	-16,79	10,98	-37,38	-4,53	-19,73	1,21
Aprile	-28,02	-6,98	-8,54	12,60	-19,78	1,58	-2,12	19,13	-32,95	-12,19	-13,56	4,90
Maggio	-39,91	-16,72	-23,26	-0,21	-39,42	-15,46	-20,21	1,02	-39,50	-16,52	-23,77	-0,92
Giugno	-42,57	-26,79	-28,00	-12,39	-39,96	-24,90	-26,37	-11,88	-40,91	-25,63	-27,43	-12,33
Luglio	-33,81	-18,06	-19,95	-4,09	-35,55	-21,69	-22,14	-9,06	-31,76	-16,19	-18,25	-2,47
Agosto	-32,37	-16,12	-16,49	0,21	-37,64	-21,71	-23,14	-8,42	-30,08	-14,90	-16,26	-0,54
Settembre	-33,96	-7,85	-11,35	16,83	-36,12	-13,77	-15,21	9,39	-32,73	-8,50	-12,16	15,57
Ottobre	-40,56	-20,13	-24,82	-7,01	-45,00	-18,99	-26,65	-2,09	-39,81	-21,36	-24,45	-7,28
Novembre	-25,75	-1,62	-5,66	18,10	-27,86	-3,29	-7,52	17,64	-25,99	-1,66	-6,13	19,23
Dicembre	-36,57	-15,27	-16,56	3,39	-37,62	-15,94	-18,71	2,23	-38,17	-16,84	-18,31	3,12

Tab. 18 Significatività delle variazioni di PM<sub>10</sub>, nelle centraline valide, per tipo di variazione e per mese (%)

mese	Δ non significativa			Δ significativa		Totale significativa
	Δ negativa	Δ positiva	Totale non significativa	Δ negativa	Δ positiva	
<b>Marzo</b>	17,76%	9,42%	27,18%	54,42%	18,41%	72,82%
<b>Aprile</b>	22,15%	17,17%	39,32%	40,81%	19,87%	60,68%
<b>Maggio</b>	20,67%	11,40%	32,08%	54,73%	13,19%	67,92%
<b>Giugno</b>	23,00%	8,20%	31,20%	65,18%	3,62%	68,80%
<b>Luglio</b>	25,22%	12,75%	37,97%	54,29%	7,74%	62,03%
<b>Agosto</b>	26,21%	15,61%	41,82%	49,41%	8,77%	58,18%
<b>Settembre</b>	17,32%	13,83%	31,14%	43,61%	25,25%	68,86%
<b>Ottobre</b>	17,59%	8,73%	26,32%	63,23%	10,45%	73,68%
<b>Novembre</b>	15,71%	11,73%	27,44%	42,22%	30,34%	72,56%
<b>Dicembre</b>	18,82%	11,91%	30,73%	52,71%	16,56%	69,27%
<b>Totale</b>	<b>20,49%</b>	<b>12,06%</b>	<b>32,55%</b>	<b>52,16%</b>	<b>15,29%</b>	<b>67,45%</b>

### 3.3.6 Particolato atmosferico: PM<sub>2.5</sub>

Nelle cartine seguenti sono riportate le differenze medie percentuali tra il dato previsto dal modello e il dato osservato per le singole stazioni su cui è stato rilevato il PM<sub>2.5</sub>. Le tre mappe si riferiscono ai periodi 1 marzo – 30 aprile, 1 maggio – 30 settembre e 1 ottobre - 31 dicembre 2020.

I dati presentati si riferiscono alle sole stazioni di monitoraggio che hanno superato la fase di validazione.



Fig. 93 PM<sub>2.5</sub> Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)



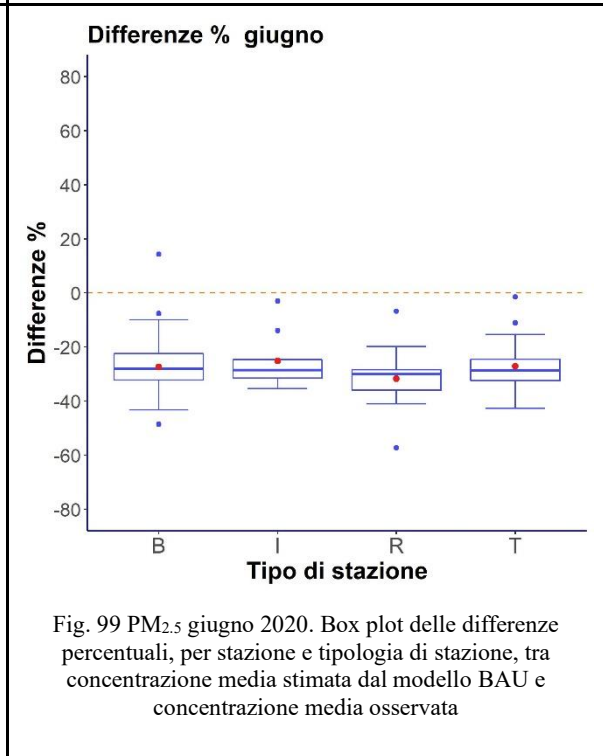
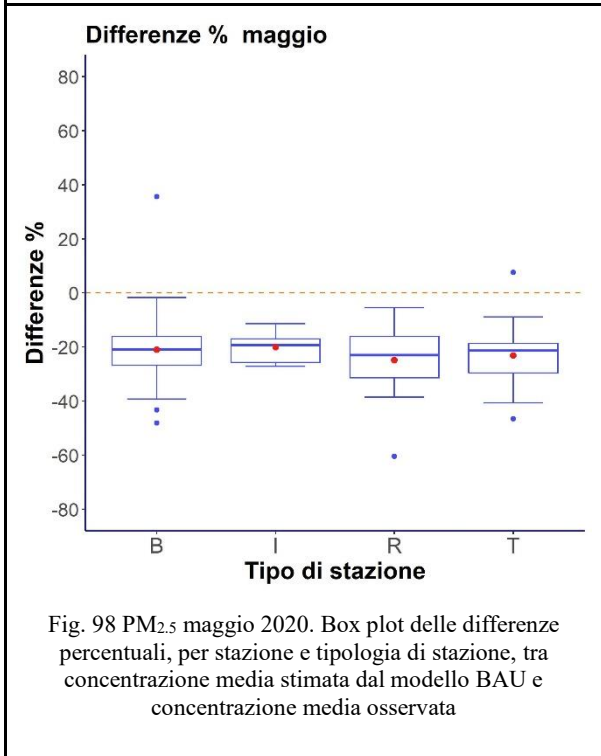
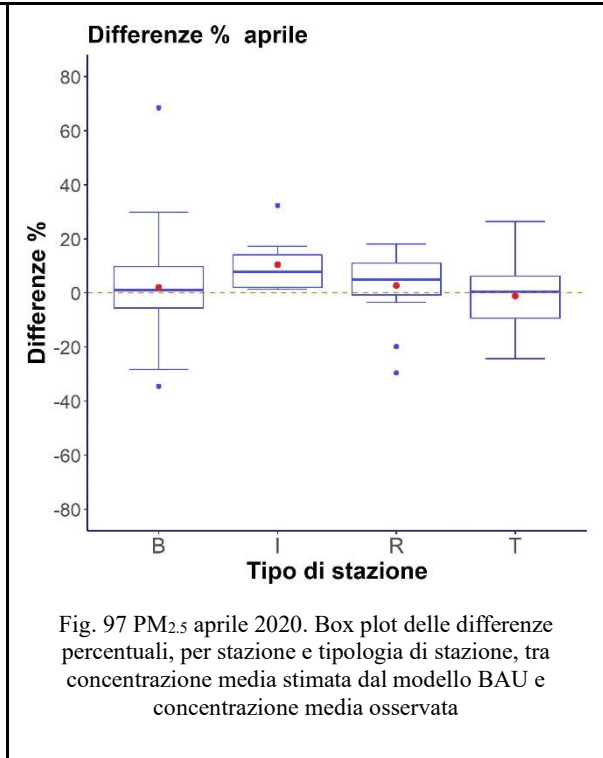
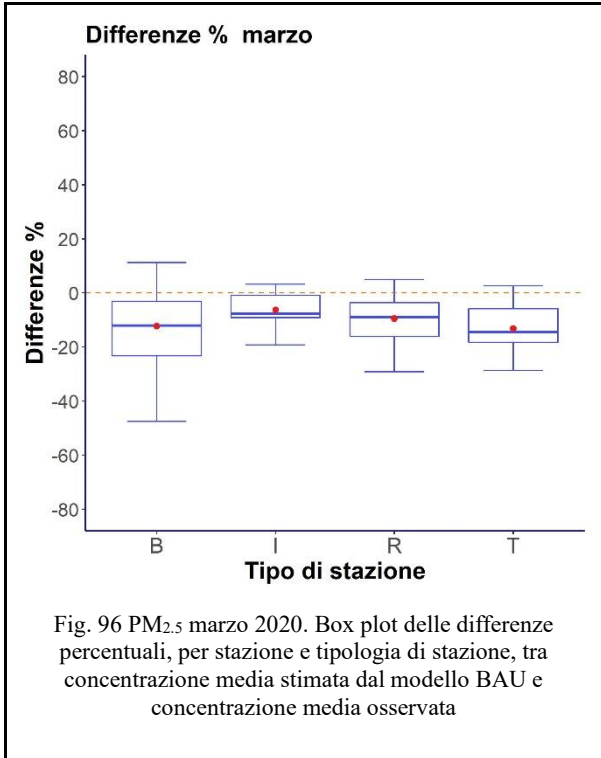
Fig. 94 PM<sub>2.5</sub> Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)

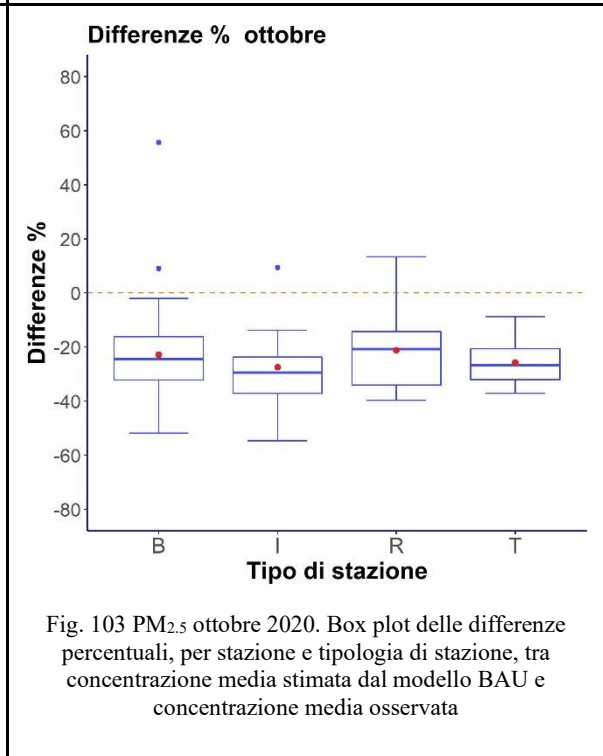
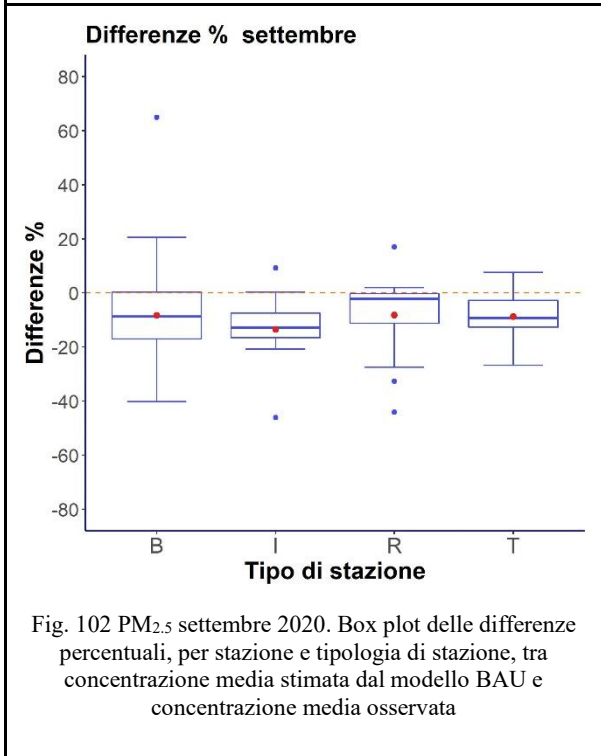
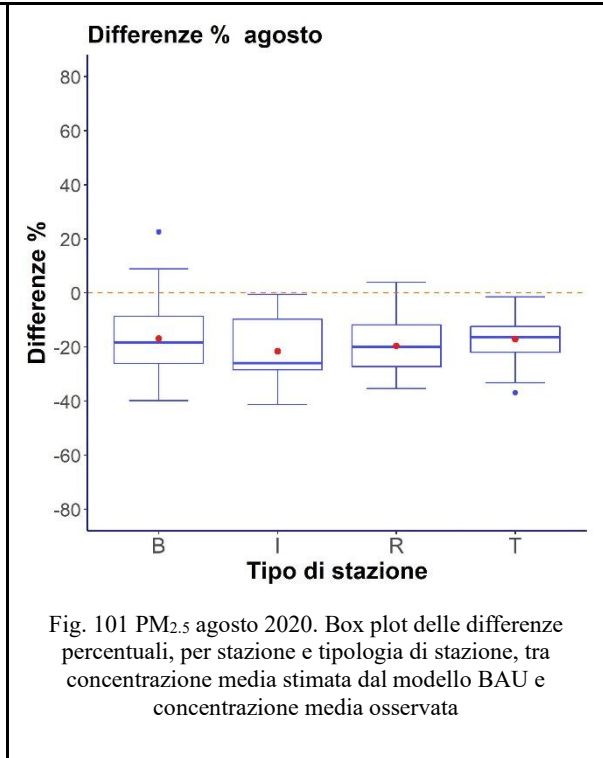
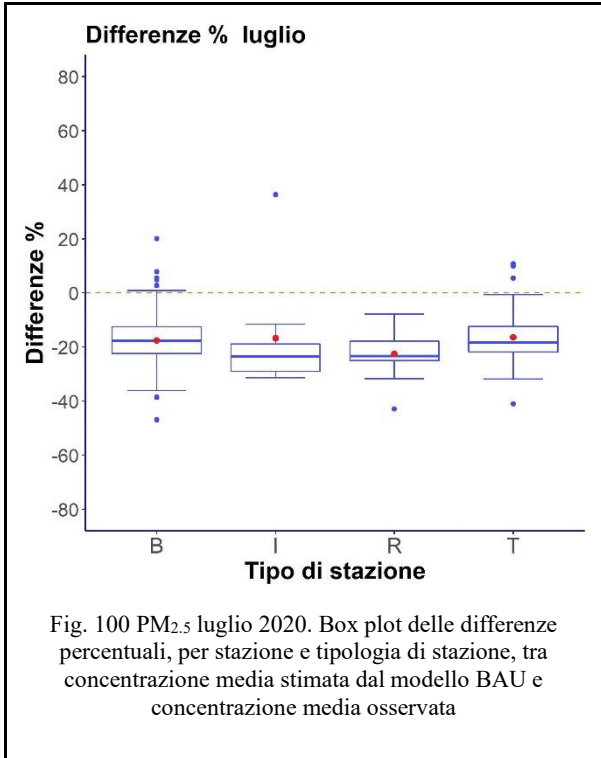


Fig. 95 PM<sub>2.5</sub> Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)

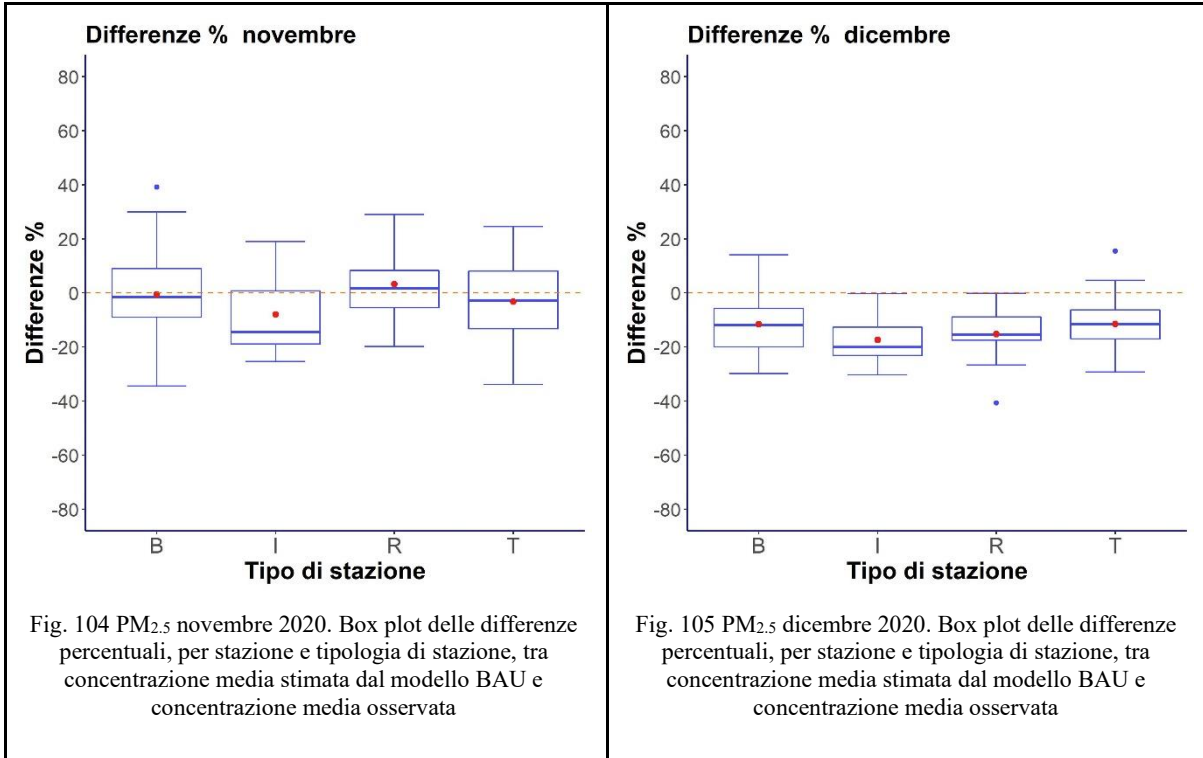
Le differenze medie annue risultano generalmente negative, con valori osservati inferiori ai valori previsti dal modello, con una certa disomogeneità tra le diverse stazioni (a differenza di quanto osservato per il PM<sub>10</sub> per cui la distribuzione spaziale risultava più uniforme). Non si evidenzia un gradiente Nord-Sud che faccia ipotizzare un diverso comportamento delle stazioni in base alla latitudine.

I boxplot seguenti rappresentano la distribuzione, mese per mese, delle differenze medie percentuali tra valore osservato e previsto per le diverse stazioni, raggruppate per tipologia di stazione.









Il comportamento del PM<sub>2.5</sub> è abbastanza simile a quello osservato per il PM<sub>10</sub>. Le differenze percentuali tra il valore osservato e previsto indicano una diminuzione delle concentrazioni nei mesi di maggio, giugno, luglio, agosto, ottobre e dicembre, in maniera abbastanza omogenea per le 4 tipologie di stazione (differenze medie pari o inferiori a -20%, mentre nel periodo di validazione erano contenute entro -10%). A differenza di quanto avveniva per il PM<sub>10</sub>, un calo delle concentrazioni si osserva anche per il mese di marzo (seppure inferiore a quello osservato per i mesi precedentemente indicati), probabilmente imputabile al fatto che questa frazione granulometrica ha risentito in maniera minore del trasporto di sabbia a lunga distanza che ha caratterizzato il mese di marzo. Non si osservano invece apprezzabili diminuzioni nelle concentrazioni del mese di aprile, che pure è stato caratterizzato da misure di contenimento per l'emergenza sanitaria molto importanti.

Tab. 19 Statistiche descrittive delle variazioni di PM<sub>2.5</sub>, nelle centraline valide, per tipo di stazione e per mese (µg/m<sup>3</sup>)

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err
Marzo	-6,42	-2,36	-2,73	1,16	1,58	-4,96	-1,37	-2,09	1,27	1,47	-6,16	-2,65	-2,80	0,58	1,51
Aprile	-3,07	0,26	-0,09	3,19	1,24	-2,36	1,25	1,04	4,54	1,17	-3,70	-0,17	-0,79	2,91	1,22
Maggio	-4,34	-2,27	-2,54	-0,46	1,16	-4,28	-2,13	-2,22	-0,53	1,14	-4,91	-2,76	-3,09	-0,97	1,15
Giugno	-4,29	-2,79	-2,81	-1,28	1,14	-4,56	-2,82	-2,77	-1,34	1,12	-4,64	-3,09	-3,15	-1,58	1,14
Luglio	-3,90	-2,08	-2,25	-0,36	1,14	-4,44	-1,95	-2,57	-0,76	1,14	-3,81	-2,07	-2,24	-0,37	1,14
Agosto	-4,14	-2,05	-2,37	-0,25	1,16	-4,86	-2,52	-2,71	-0,06	1,14	-4,43	-2,30	-2,51	-0,11	1,17
Settembre	-3,85	-0,80	-1,53	2,27	1,17	-4,02	-1,03	-1,92	1,97	1,16	-4,14	-0,98	-1,67	1,85	1,18
Ottobre	-6,02	-2,99	-3,30	-0,82	1,42	-7,19	-3,48	-3,47	-1,32	1,47	-6,01	-3,83	-3,74	-1,77	1,39
Novembre	-5,65	0,43	-1,00	5,19	1,86	-6,43	-0,36	-1,47	3,49	1,95	-6,16	0,05	-1,71	4,58	1,81
Dicembre	-6,97	-2,32	-2,63	1,82	1,89	-7,47	-3,07	-2,78	0,80	1,80	-6,07	-1,36	-2,58	1,75	1,80

Tab. 20 Statistiche descrittive delle variazioni di PM<sub>2.5</sub>, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-37,88	-11,75	-18,75	7,81	-34,83	-6,72	-16,21	11,18	-35,04	-13,17	-17,95	3,96
Aprile	-25,98	2,03	-0,71	26,28	-19,36	11,26	8,18	36,78	-27,80	-0,98	-6,44	21,24
Maggio	-43,13	-21,70	-26,40	-4,76	-37,58	-19,41	-22,66	-5,93	-42,26	-23,18	-28,24	-9,46
Giugno	-44,14	-27,93	-29,95	-13,34	-43,93	-26,39	-27,90	-13,48	-41,76	-27,28	-29,12	-14,64
Luglio	-35,12	-18,32	-19,62	-3,52	-36,79	-17,61	-22,54	-6,90	-31,65	-16,46	-17,60	-3,29
Agosto	-36,36	-17,54	-20,30	-2,16	-39,79	-20,44	-22,94	-0,53	-35,05	-17,11	-20,50	-0,75
Settembre	-37,96	-8,21	-13,94	18,79	-42,53	-10,97	-17,56	15,41	-35,20	-8,79	-13,43	13,60
Ottobre	-46,12	-22,73	-27,98	-6,85	-48,62	-27,13	-34,37	-15,51	-41,13	-25,78	-29,00	-13,89
Novembre	-26,82	-0,12	-4,79	20,96	-34,80	-5,41	-9,75	15,29	-27,86	-3,21	-7,31	16,65
Dicembre	-35,15	-12,21	-14,31	8,37	-37,36	-17,17	-21,54	4,21	-34,73	-11,43	-14,74	8,26

Tab. 21 Significatività delle variazioni di PM<sub>2,5</sub>, nelle centraline valide, per tipo di variazione e per mese (%)

mese	Δ non significativa			Δ significativa		Totale significativa
	Δ negativa	Δ positiva	Totale non significativa	Δ negativa	Δ positiva	
<b>Marzo</b>	18,81%	11,55%	30,36%	51,01%	18,63%	69,64%
<b>Aprile</b>	19,77%	16,34%	36,11%	32,29%	31,60%	63,89%
<b>Maggio</b>	22,19%	11,80%	34,00%	57,45%	8,56%	66,00%
<b>Giugno</b>	24,28%	9,19%	33,47%	63,70%	2,83%	66,53%
<b>Luglio</b>	27,17%	13,68%	40,85%	52,44%	6,71%	59,15%
<b>Agosto</b>	23,63%	13,18%	36,82%	52,91%	10,27%	63,18%
<b>Settembre</b>	18,23%	12,33%	30,56%	44,38%	25,06%	69,44%
<b>Ottobre</b>	18,44%	8,21%	26,64%	64,38%	8,97%	73,36%
<b>Novembre</b>	16,06%	13,21%	29,27%	40,55%	30,17%	70,73%
<b>Dicembre</b>	18,39%	12,62%	31,01%	48,71%	20,28%	68,99%
<b>Totale</b>	<b>20,72%</b>	<b>12,21%</b>	<b>32,92%</b>	<b>50,83%</b>	<b>16,24%</b>	<b>67,08%</b>



## 4 CONCLUSIONI

Le misure restrittive che hanno avuto luogo nel 2020 per cercare di contenere l'epidemia di SARS-CoV-2 hanno dato luogo ad una diminuzione di alcune delle attività antropiche responsabili dell'inquinamento atmosferico (principalmente traffico automobilistico, aereo e navale ed attività produttive e commerciali). Si è pertanto voluto valutare se c'è stata, e in quale misura, una corrispondente diminuzione delle concentrazioni in atmosfera dei principali inquinanti rilevati. A tal fine è stato necessario controllare l'effetto delle variabili meteorologiche, che come noto, hanno un impatto sulla concentrazione degli inquinanti spesso predominante rispetto all'impatto imputabile alla variazione delle emissioni.

Lo studio qui presentato ha utilizzato a tal scopo il modello Random Forest, che ha permesso di stimare le concentrazioni di inquinanti che si sarebbero osservate in atmosfera con le condizioni meteorologiche effettivamente verificatesi nel 2020, ma in assenza delle misure restrittive messe in atto.

Il modello è stato addestrato sugli anni dal 01/01/2013 al 31/12/2019; il periodo dal 1° gennaio 2020 al 14 febbraio 2020 - durante il quale non è stata messa in atto alcuna misura di contenimento - è stato utilizzato come periodo di validazione del modello e tutti i confronti tra valore osservato e previsto che sono stati effettuati nei mesi successivi sono stati valutati tenendo presente l'incertezza e l'errore del modello osservato durante il periodo di validazione.

Il periodo di lockdown che ha caratterizzato i mesi di marzo, aprile e maggio del 2020 ha visto una riduzione importante delle concentrazioni di NO<sub>2</sub> (differenze percentuali medie di -8%, -37% e -50% rispettivamente per i mesi di marzo, aprile e maggio nelle stazioni di traffico, a fronte di una media delle differenze vicina al 5% osservata durante il periodo di validazione) ed una riduzione di CO e Benzene maggiore di quella osservata durante il periodo di validazione solo per alcune tipologie di stazioni. Non si osserva alcuna variazione significativa delle concentrazioni di O<sub>3</sub>. Per quel che riguarda il particolato, nel mese di maggio si osserva una diminuzione sia di PM<sub>10</sub> che di PM<sub>2,5</sub>, mentre nel mese di marzo si osserva un calo solo delle concentrazioni di PM<sub>2,5</sub>, ma il PM<sub>10</sub> ha risentito di un trasporto di polveri a larga scala che il modello non era in grado di prevedere e questo può avere



in parte compensato l'eventuale calo di  $PM_{10}$  di origine antropica. Il mese di aprile non mostra invece una riduzione apprezzabile per alcuno dei due tagli dimensionali.

I mesi estivi (giugno, luglio e agosto) sono stati caratterizzati da una diminuzione di  $NO_2$  (seppure inferiore a quella osservata durante il lockdown primaverile): media delle differenze vicina a -20% per le 4 tipologie di stazioni. Il benzene ha visto una diminuzione ad eccezione delle stazioni di tipo industriale e molto marcata (differenze medie fino a -60% circa) per le stazioni rurali, mentre per CO e  $O_3$  non si osserva una diminuzione apprezzabile. Il particolato mostra differenze percentuali medie intorno a -20% circa in entrambe le frazioni granulometriche per le 4 tipologie di stazioni, mettendo in evidenza un calo maggiore rispetto a quello osservato durante il periodo di lockdown primaverile. Per quel che riguarda i mesi successivi, da settembre a dicembre, il benzene presenta, nei mesi di settembre e ottobre differenze maggiori, in valore assoluto, a quelle osservate durante il periodo di validazione, con molta variabilità tra le diverse tipologie di stazioni, mentre non sembrano evidenti diminuzioni della concentrazione di CO. L' $NO_2$  presenta un calo delle concentrazioni anche durante questi mesi, seppure inferiore a quello osservato durante il lockdown primaverile, con differenze medie che si attestano fra i -5% e -20% circa nei vari mesi e nelle diverse tipologie di stazioni. L' $O_3$  non sembra mostrare differenze apprezzabili nemmeno durante questi mesi autunnali-invernali, mentre il particolato mostra un calo delle concentrazioni nei due tagli dimensionali ( $PM_{10}$  e  $PM_{2.5}$ ) nei mesi di ottobre e dicembre.

L'inquinante che ha mostrato le diminuzioni più consistenti è quindi l' $NO_2$ , che oltre a mostrare un calo maggiore rispetto agli altri inquinanti durante il periodo di lockdown, mostra una diminuzione in tutti i mesi da marzo in poi, a indicare un effetto delle misure di contenimento anche successive ai mesi di lockdown primaverile (riguardanti principalmente gli spostamenti). L' $O_3$  non mostra invece alcuna diminuzione durante tutto l'anno, probabilmente a causa della sua natura secondaria principalmente influenzata dalle dinamiche di formazione fotochimica.

Per quel che riguarda i due gas primari (benzene e CO), la diminuzione si osserva in maniera discontinua tra i diversi mesi e nelle diverse tipologie di stazioni e rimane quindi argomento da approfondire.



Il particolato presenta un andamento assolutamente paragonabile nei due tagli dimensionali, fatta eccezione per il mese di marzo a causa del trasporto di polveri a larga scala che ha impattato evidentemente in maniera maggiore sulla frazione più grossolana. Per tale inquinante si osservano diminuzioni non importanti come quelle riscontrate per l'NO<sub>2</sub> ma comunque apprezzabili, abbastanza simili tra le diverse tipologie di stazioni, ma discontinue nei mesi, con un calo estivo inaspettatamente più importante rispetto a quello primaverile. Sarà interessante indagare l'impatto, sulle concentrazioni di questo inquinante, della combustione di legna, che durante i mesi estivi si abbatte, mettendo forse in maggior risalto l'effetto delle misure di contenimento relative agli spostamenti (come ad esempio la possibilità di lavorare in smart working).



## 5 BIBLIOGRAFIA

- Araki, Shin , Masayuki Shima, e Kouhei Yamamoto. «patiotemporal land use random forest model for estimating metropolitan no2 exposure in japan.» *Science of The Total Environment*, 2018: 634: 1269 – 1277.
- Breiman, Leo. «Bagging predictors machine learning.» *Machine learning*, 1996.
- Breiman, Leo. «Random forests.» *Machine Learning*, 2001: 45(1):5–32.
- Breiman, Leo, J. H. Friedman, e R. A. Olshen. *Classification and Regression Trees*. New York: Routledge, 1984.
- Coulston, John W., Christine E. Blinn, Valerie A. Thomas, e Randolph H. Wynne. «Approximating prediction uncertainty for random forest regression models.» *Photogrammetric Engineering & Remote Sensing*, 2016: 189-197.
- Genuer, Robin, Jean-Michel Poggi, e Christine Tuleau-Malot. «Variable selection using random forests.» *Pattern Recognition Letters*, 2010: 31:2225–2236.
- Grange, Stuart K, James D. Lee, Will S Drysdale, Alastair C Lewis, Christoph Hueglin, e David C Carslaw. «COVID-19 lockdowns highlight a risk of increasing ozone pollution in European urban areas.» *Atmospheric Chemistry and Physics*, 2021: 4169-4185.
- Grange, Stuart K, e David C Carslaw. «Using meteorological normalisation to detect interventions in air quality time series.» *Science of the Total Environment*, 2019: 653, 578-588.
- Grange, Stuart K., David Carslaw, Alastair Lewis, Eirini Boleti, e Christoph Hueglin. «andom forest meteorological normalisation models for Swiss PM 10 trend analysis.» *Atmospheric Chemistry and Physics*, 2018: 18(9), 6223-6239.
- Meinshausen, Nicolai, e Greg Ridgeway. «Quantile regression forests.» *Journal of machine learning research*, 2006.
- Mentch, Lucas, e Giles Hooker. «Quantifying uncertainty in random forests via confidence intervals and hypothesis tests.» *The Journal of Machine Learning Research*, 2016: 841-881.
- Petetin, Hervé, et al. «Meteorology-normalized impact of the COVID-19 lockdown upon NO 2 pollution in Spain.» *Atmospheric Chemistry and Physics*, 2020: 20(18), 11119-11141.



- R Core Team. «R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.» 2021.
- Rahman, Md Mahmudur, Jayanandana Karunasinghe, Sam Clifford, Luke D Knibbs, e Lidia Morawska. «New insights into the spatial distribution of particle number concentrations by applying non-parametric land use regression modelling.» *Science of The Total Environment*, 2020: 702:134708.
- Stafoggia, Massimo, et al. «Estimation of daily pm10 and pm2.5 concentrations in italy, 2013–2015, using a spatiotemporal land-use random-forest model.» *Environment International*, 2019: 124:170 – 179.
- Wager, Stefan, Trevor Hastie, e Bradley Efron. «Confidence intervals for random forests: The jackknife and the infinitesimal jackknife.» *The Journal of Machine Learning Research*, 2014: 1625-1651.
- Wright MN, Ziegler A. «ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.» *Journal of Statistical Software*, 2017: 77(1), 1–17.





## **APPENDICE A OZONO: MASSIMO MEDIA MOBILE DI 8 H.**

Nelle cartine seguenti sono riportate le differenze medie percentuali tra il dato previsto dal modello e il dato osservato per le singole stazioni su cui è stato rilevato l'ozono.

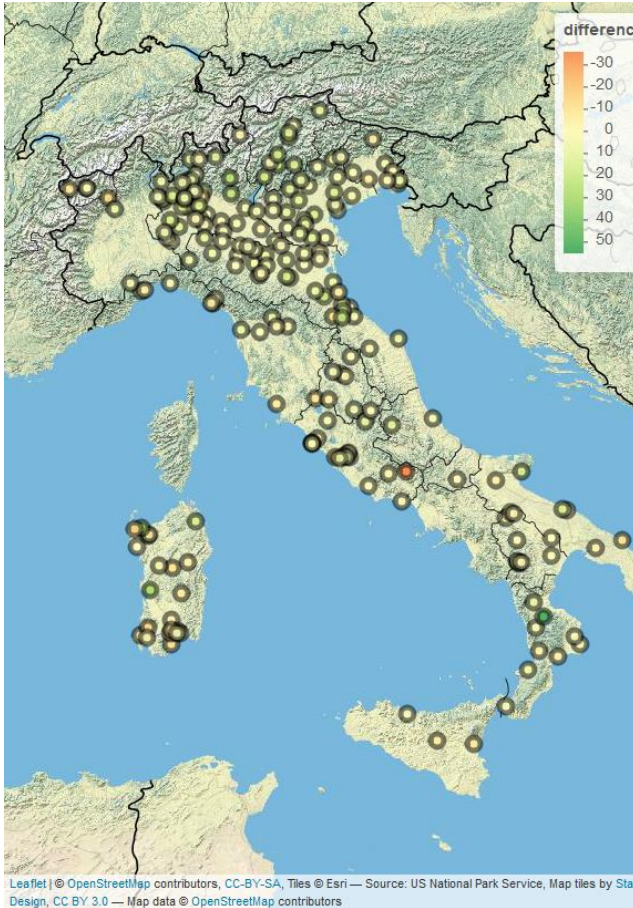


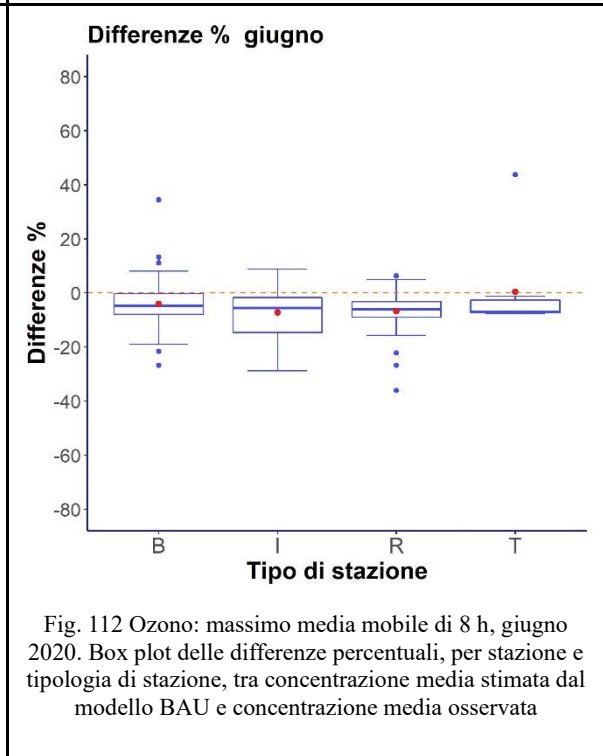
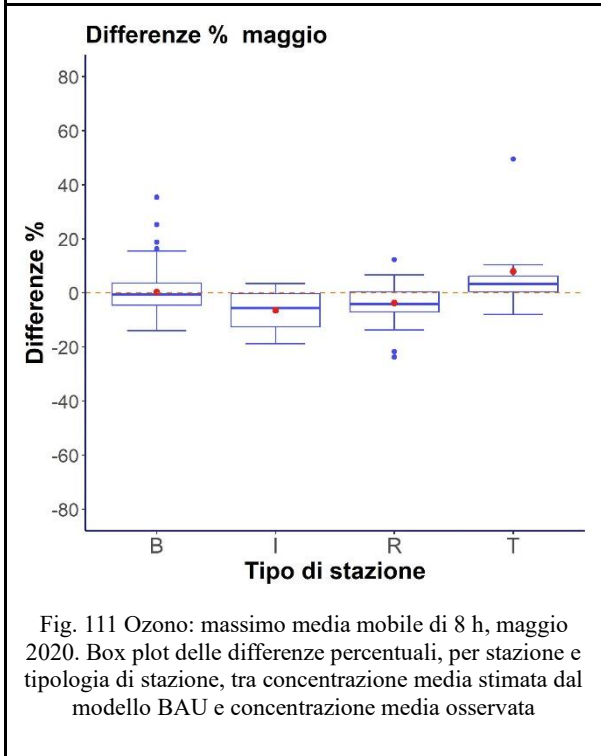
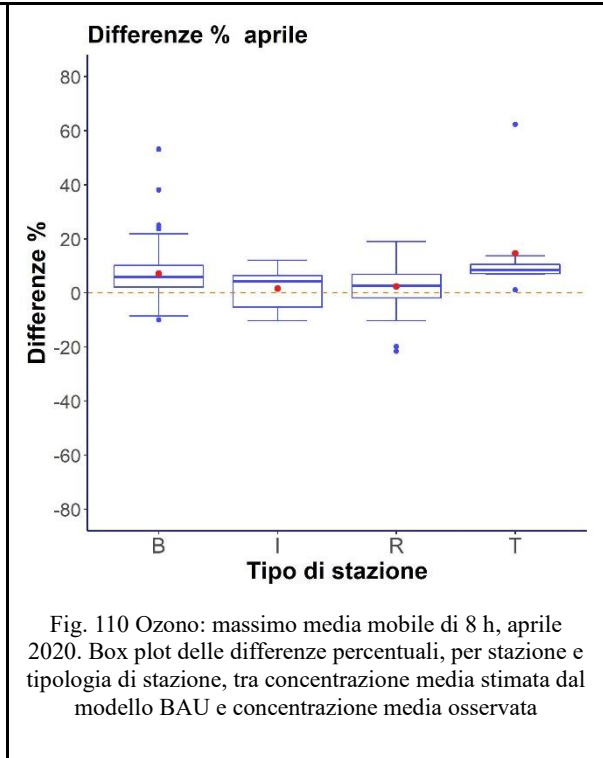
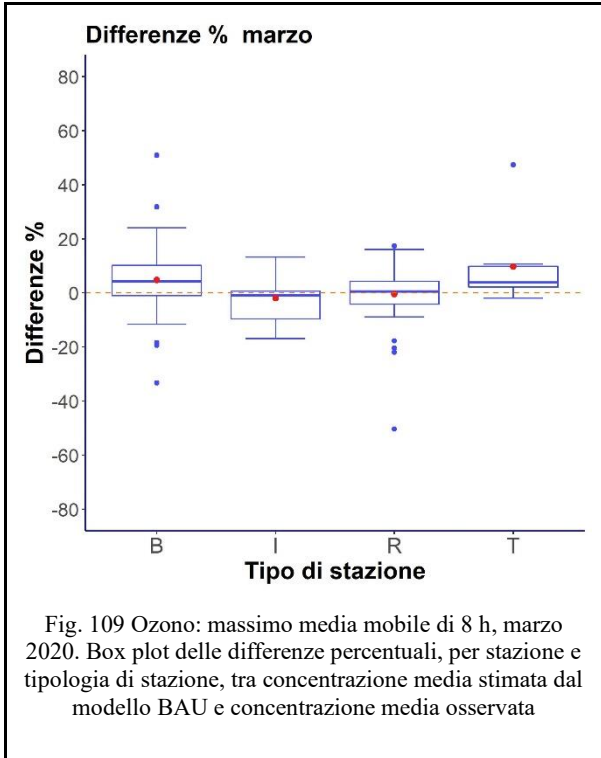
Fig. 106 Ozono: massimo media mobile di 8 h Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 marzo al 30 aprile 2020)

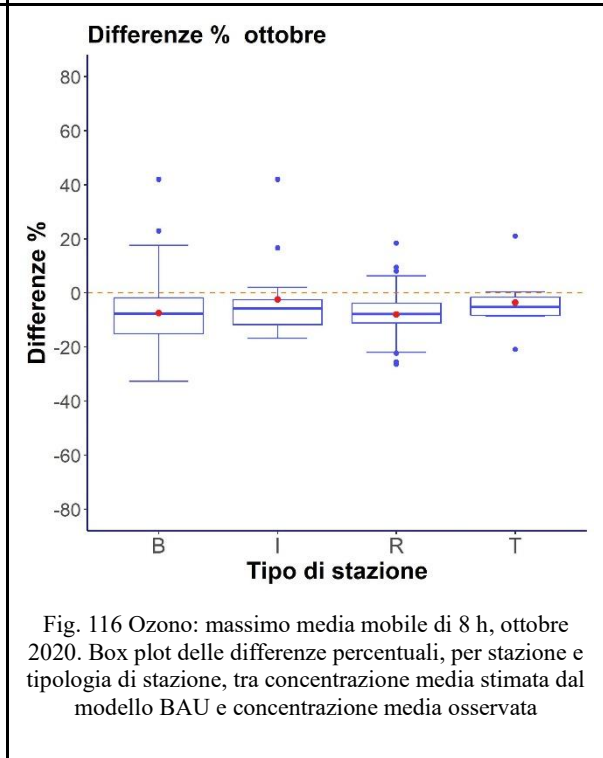
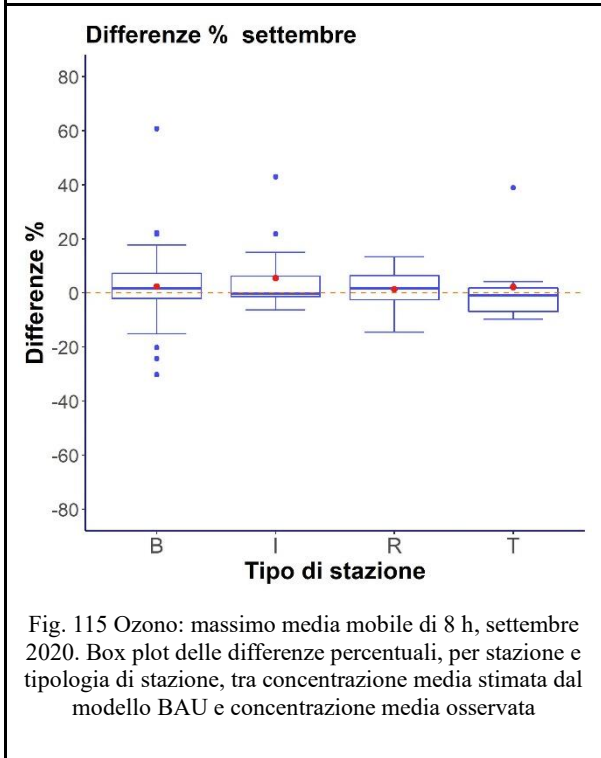
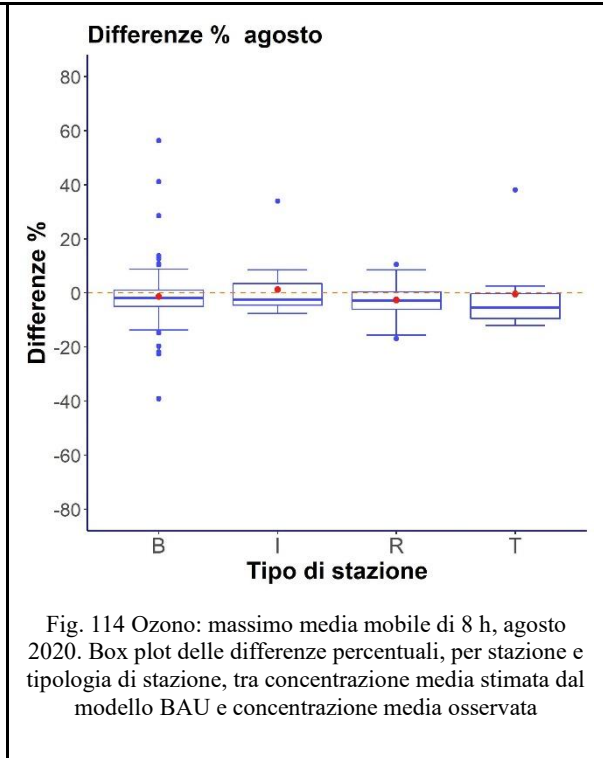
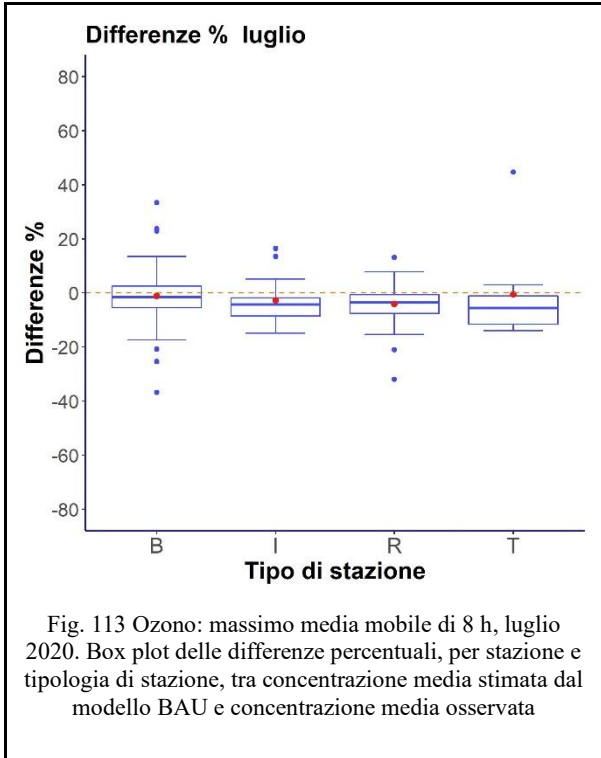


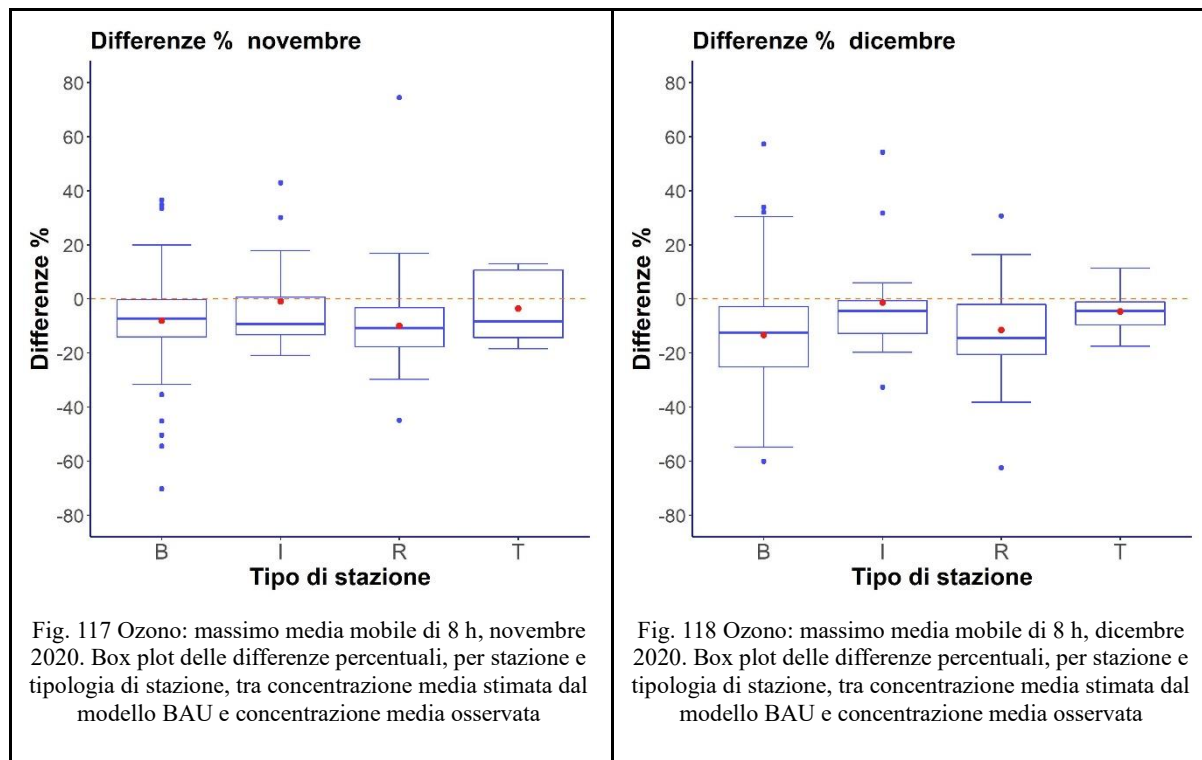
Fig. 107 Ozono: massimo media mobile di 8 h Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 maggio al 30 settembre 2020)



Fig. 108 Ozono: massimo media mobile di 8 h Mappa delle differenze percentuali tra concentrazione media stimata dal modello BAU e concentrazione media osservata (periodo dal 1 ottobre al 31 dicembre 2020)







Tab. 22 Statistiche descrittive delle variazioni di Ozono, nelle centraline valide, per tipo di stazione e per mese ( $\mu\text{g}/\text{m}^3$ )

Mese	Fondo					Industriale					Traffico				
	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err	25th	Media	Mediana	75th	Std. err
Marzo	-6,06	1,19	0,64	8,42	2,69	-7,59	-0,78	-1,24	5,22	2,21	-2,80	5,14	3,49	10,13	2,11
Aprile	-3,28	5,03	4,48	12,88	2,68	-6,49	2,66	1,54	10,65	2,21	-1,57	9,53	7,39	17,77	2,15
Maggio	-9,29	-1,07	-1,53	6,66	2,66	-11,89	-4,88	-4,86	2,80	2,21	-4,18	4,17	3,01	10,91	2,09
Giugno	-13,96	-5,22	-5,21	3,21	2,78	-13,72	-6,11	-6,48	1,63	2,19	-11,02	-2,02	-3,37	3,85	2,23
Luglio	-12,30	-2,17	-3,37	6,81	2,93	-12,61	-2,85	-4,02	5,62	2,30	-12,28	-1,95	-3,85	6,47	2,28
Agosto	-10,81	-1,78	-2,54	6,32	2,87	-9,83	-1,18	-2,30	6,80	2,19	-11,82	-1,09	-2,93	6,92	2,29
Settembre	-7,14	1,96	1,17	10,27	2,76	-4,96	3,80	3,08	11,33	2,23	-7,38	1,11	-0,83	8,77	2,22
Ottobre	-12,14	-4,07	-4,85	3,47	2,66	-11,88	-3,11	-4,46	4,05	2,08	-10,11	-2,20	-2,95	5,06	2,13
Novembre	-9,45	-3,15	-3,18	3,15	2,57	-11,73	-3,54	-5,31	3,55	2,09	-8,72	-2,24	-2,35	4,68	2,16
Dicembre	-9,49	-3,17	-3,99	3,32	2,58	-9,26	-1,85	-2,52	4,08	2,08	-7,18	-0,92	-1,45	5,49	2,05

Tab. 23 Statistiche descrittive delle variazioni di Ozono, nelle centraline valide, per tipo di stazione e per mese (%)

Mese	Fondo				Industriale				Traffico			
	25th	Media	Mediana	75th	25th	Media	Mediana	75th	25th	Media	Mediana	75th
Marzo	-7,72	3,32	0,84	11,62	-8,83	-0,46	-1,39	6,29	-3,45	9,67	4,88	14,67
Aprile	-3,50	5,74	4,64	13,67	-7,38	2,34	1,60	11,06	-1,84	14,50	8,87	20,46
Maggio	-9,69	-0,85	-1,62	7,03	-12,35	-5,22	-5,22	2,82	-4,30	8,05	3,65	12,67
Giugno	-13,96	-4,84	-5,23	3,26	-14,91	-6,32	-6,24	1,53	-12,20	0,17	-3,76	4,37
Luglio	-10,89	-2,09	-3,08	6,07	-11,65	-2,62	-3,76	5,34	-13,67	-0,49	-4,44	7,06
Agosto	-10,43	-1,66	-2,43	5,91	-9,31	-0,25	-2,25	6,60	-14,02	-0,31	-3,28	8,87
Settembre	-8,38	1,97	1,31	11,26	-5,34	4,96	3,14	11,60	-9,58	2,26	-1,00	10,51
Ottobre	-20,88	-7,44	-7,90	5,76	-15,07	-3,38	-5,69	4,72	-15,82	-3,63	-4,39	8,35
Novembre	-27,21	-8,20	-8,70	8,46	-18,47	-4,65	-8,22	6,38	-21,34	-3,38	-6,69	10,22
Dicembre	-40,68	-13,09	-11,75	7,95	-16,12	-4,08	-4,32	6,21	-18,85	-4,66	-2,82	10,87

Tab. 24 Significatività delle variazioni di Ozono, nelle centraline valide, per tipo di variazione e per mese (%)

mese	Δ non significativa			Δ significativa		
	Δ negativa	Δ positiva	Totale non significativa	Δ negativa	Δ positiva	Totale significativa
Marzo	18,70%	18,37%	37,07%	29,11%	33,82%	62,93%
Aprile	14,13%	17,18%	31,31%	21,75%	46,94%	68,69%
Maggio	17,04%	15,10%	32,14%	39,16%	28,70%	67,86%
Giugno	15,81%	13,68%	29,49%	50,54%	19,97%	70,51%
Luglio	15,32%	12,55%	27,88%	44,72%	27,41%	72,12%
Agosto	16,53%	14,74%	31,27%	41,50%	27,23%	68,73%
Settembre	15,09%	15,80%	30,89%	30,61%	38,50%	69,11%
Ottobre	16,03%	13,22%	29,25%	49,08%	21,68%	70,75%
Novembre	20,49%	14,69%	35,18%	43,75%	21,06%	64,82%
Dicembre	19,76%	13,48%	33,24%	44,72%	22,04%	66,76%
<b>Totale</b>	<b>16,88%</b>	<b>14,90%</b>	<b>31,78%</b>	<b>39,44%</b>	<b>28,78%</b>	<b>68,22%</b>