



Interreg



UNIONE EUROPEA

MARITTIMO-IT FR-MARITIME



GIREPAM

Progetto GIREPAM

Gestione Integrata delle Reti Ecologiche attraverso i Parchi e le Aree Marine

Prodotto T1.3.5

Azioni pilota per ambiti dunali e zone di transizione (foci fluviali, lagune costiere)

Marzo 2020



Genova, 30 marzo 2020

*La cooperazione al cuore del Mediterraneo
La coopération au cœur de la Méditerranée*

Indice generale

Studio sulla capacità di carico in un sito sensibile dell'AMP Portofino (S. Fruttuoso di Camogli) in convenzione con UNIGE.....	3
1 Realizzazione di un modello per la valutazione dei flussi turistici nella baia di San Fruttuoso di Camogli.....	3
1.1 <i>Metodi utilizzati</i>	3
1.2 <i>Risultati</i>	5
2 Bibliografia	11

Studio sulla capacità di carico in un sito sensibile dell'AMP Portofino (S. Fruttuoso di Camogli) in convenzione con UNIGE

Nell'ambito di questo prodotto è prevista la realizzazione di un modello per la valutazione dei flussi turistici nella baia di San Fruttuoso di Camogli.

1 Realizzazione di un modello per la valutazione dei flussi turistici nella baia di San Fruttuoso di Camogli

1.1 Metodi utilizzati

Il modello per la valutazione dei flussi turistici nella baia di San Fruttuoso ha permesso di individuare le principali forzanti che guidano gli arrivi e le presenze nella baia e consente di predire i flussi turistici attesi in relazione alla stagione e alle principali condizioni meteo-climatiche. Il modello proposto è realizzato al fine di fornire al gestore uno strumento previsionale per la regolazione degli accessi alla baia anche nell'ottica di non superare la capacità di carico individuata.

A tale scopo, nel 2018 sono stati realizzati dei monitoraggi, al fine di raccogliere le informazioni di base per la realizzazione dello studio. In particolare, sono state realizzate 30 giornate di campionamento nei mesi tra giugno e novembre. Il campionamento è stato realizzato tramite operatori opportunamente formati che stazionavano presso la baia nella fascia oraria compresa tra le 9 e le 19.

Gli operatori provvedevano ad annotare su apposite schede gli arrivi e le partenze dei fruitori suddividendole, per ogni fascia oraria, tra arrivi/partenze tramite battelli di linea e tramite mezzi nautici privati o taxi boat. Gli operatori annotavano, inoltre, per ogni fascia oraria, il numero di persone stazionanti in spiaggia e, ad inizio campionamento, le condizioni meteo-climatiche. A tale scopo sono state individuate, per le condizioni del cielo e del mare, delle classi principali riportate in Tabella 1.

Legenda	Cielo	Mare
1	sereno	calmo
2	poco nuvoloso	quasi calmo/poco mosso
3	nuvoloso	mosso
4	pioggia	molto mosso

Tabella 1. Classificazione delle condizioni del cielo e del mare adottata

Al fine del monitoraggio sono state realizzate e fornite apposite schede in formato sia cartaceo sia digitale.

Per quanto concerne gli arrivi tramite la sentieristica, è presente, nella baia, un contapassi gestito dal Parco Naturale Regionale di Portofino che ha provveduto a fornire i dati di arrivi e partenze per le giornate campionate.

Il modello è stato sviluppato con la metodologia di regressione Random Forest (RF) che garantisce elevate capacità di accoppiamento fra le variabili esplicative (i parametri del modello) e le risposte attese, nel caso specifico gli arrivi e le presenze nella baia di San Fruttuoso.

La Random Forest (RF) si basa sulla metodologia degli alberi di regressione, in grado di ottenere i valori di una variabile di risposta partendo da un certo numero di variabili esplicative, suddividendo l'insieme di dati forniti al sistema in sottogruppi. Le partizioni in sottogruppi ricorsive si basano su regole decisionali che permettono la divisione di ogni parte in porzioni di dati più piccoli.

Questo può essere rappresentato come un albero binario, una struttura gerarchica formata da nodi e spigoli, quest'ultima rappresenta un flusso di informazioni tra nodi adiacenti (Figura 1).

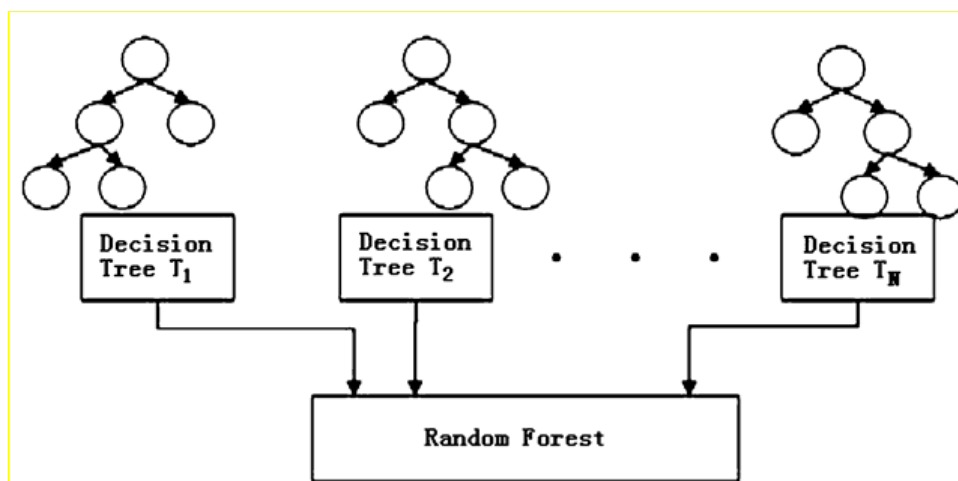


Figura 1 Esempio di come si sviluppano gli alberi decisionali

La RF è una tecnica di classificazione che accoppia la generazione di alberi decisionali alla teoria delle reti neurali (Breiman, 2001). Il principio fondamentale su cui si basa questa tecnica è uno degli elementi fondamentali di tutto il machine learning o statistical learning (Hastie et al., 2001), cioè che è più facile calibrare un modello stabilendo molte regole deboli, piuttosto che una sola regola forte.

Per far questo, la RF non sviluppa un singolo albero decisionale, ma un'intera foresta di alberi.

Ciò si ottiene in due modi: (1) una selezione casuale di variabili esplicative è scelta per far crescere ogni albero e (2) ogni albero è basato su un diverso sottoinsieme di dati casuali, creato dal processo di "bootstrapping" (Efron, 1979). Infine l'ottimale "splitting" in confronto con i dati reali è identificato e selezionato come predittore.

La porzione di dati utilizzata per la formazione dei sottoinsiemi è conosciuta come dati "in-bag", mentre i restanti si chiamano "out-of-bag". Questi ultimi non sono utilizzati per costruire l'albero, ma per fornire stime di errori di generalizzazione (Breiman, 2001). L'errore quadratico medio calcolato dalla previsione con il set di dati di test in media su tutti gli alberi è chiamato l'errore "out-

of-bag". Più la dimensione della foresta aumenta, più questo errore di generalizzazione converge sempre (Breiman, 2001).

Il numero di alberi deve essere quindi impostato sufficientemente elevato (in questo caso 1500). Gli alberi di decisione sono dotati di campioni casuali di dati per effettuare la divisione in sottoinsiemi casuali dello spazio variabile e quindi sono utilizzati per prevedere la distribuzione su tutto un insieme di dati (Kehoe et al., 2012). Successivamente, sono state identificate le variabili più rilevanti.

In particolare, l'importanza di ogni variabile esplicativa è riportata come la variazione di errore quadratico medio che si avrebbe se non se ne tenesse conto.

Dopo aver identificato le variabili più rilevanti, il successivo passo è tentare di comprendere la natura della dipendenza della variabile di risposta per ogni variabile esplicativa. La dipendenza parziale (Hastie et al., 2001) può essere utilizzata per caratterizzare graficamente le relazioni tra le singole variabili esplicative e la probabilità di presenza ottenute da RF.

1.2 Risultati

Tutti i dati raccolti tramite il campionamento effettuato sono stati digitalizzati ed inseriti in un'unica matrice, propedeutica alla realizzazione del prodotto. La matrice contiene alcuni campi fondamentali per la realizzazione del modello ovvero:

- Data
- Mese
- Stagione
- Tipologia giorno (weekend o infrasettimanale)
- Condizioni del cielo
- Condizioni del mare
- Intensità del vento
- Temperatura
- Direzione vento
- Fascia oraria
- Arrivi
- Partenze
- Arrivi via mare
- Arrivi via terra
- Partenze via mare
- Partenze via terra
- Presenze In Spiaggia

In Figura 2 è mostrata la matrice creata e che è stata impiegata nelle fasi di sviluppo e verifica del modello.

1	DATA	MESE	Stagione	Tipolo	cielo	mare	Vento	Temperatura	Direz	FASCIA ORAP	ARRIVI	PARTENZE	ARRIVI MARE	ARRIVI TER	PARTENZE	PARTENZE	Presenze in spiaggia		
23	11/06/2017	6A	W	1	1	8	24 O	10,00	87	16	37	2	5	10	103	37	7	10	54
24	11/06/2017	6A	W	1	1	8	24 O	11,00	418	9	51	0	0	19	427	51	0	19	161
25	11/06/2017	6A	W	1	1	8	24 O	12,00	212	9	61	217	5	17	221	61	222	17	185
26	11/06/2017	6A	W	1	1	8	24 O	13,00	151	28	69	212	9	13	179	69	221	13	216
27	11/06/2017	6A	W	1	1	8	24 O	14,00	143	22	37	85	12	27	165	37	97	27	187
28	11/06/2017	6A	W	1	1	8	24 O	15,00	235	14	41	130	28	25	249	41	158	25	189
29	11/06/2017	6A	W	1	1	8	24 O	16,00	59	7	5	431	23	19	66	5	454	19	190
30	11/06/2017	6A	W	1	1	8	24 O	17,00	33	6	7	289	19	6	38	7	308	6	98
31	11/06/2017	6A	W	1	1	8	24 O	18,00	4	0	1	216	7	4	4	1	223	4	84
32	11/06/2017	6A	W	1	1	8	24 O	19,00	0	0	3	0	0	0	3	0	0	0	0
33	13/06/2017	6A	L	1	1	6	24 O	10,00	0	0	6	0	0	8	0	6	0	8	29
34	13/06/2017	6A	L	1	1	6	24 O	11,00	138	16	35	57	11	13	154	35	68	13	68
35	13/06/2017	6A	L	1	1	6	24 O	12,00	158	18	17	30	19	11	176	17	49	11	95
36	13/06/2017	6A	L	1	1	6	24 O	13,00	135	18	20	48	14	9	153	20	62	9	115
37	13/06/2017	6A	L	1	1	6	24 O	14,00	47	12	29	59	11	29	59	29	70	29	93
38	13/06/2017	6A	L	1	1	6	24 O	15,00	147	7	14	91	8	12	154	14	99	12	108
39	13/06/2017	6A	L	1	1	6	24 O	16,00	52	2	7	224	11	13	54	7	235	13	128
40	13/06/2017	6A	L	1	1	6	24 O	17,00	16	12	0	177	6	2	28	0	183	2	54
41	13/06/2017	6A	L	1	1	6	24 O	18,00	0	0	2	74	0	3	0	2	74	3	0
42	13/06/2017	6A	L	1	1	6	24 O	19,00	0	0	0	0	0	1	0	0	0	1	0
43	16/06/2017	6A	L	2	1	6	25 O	10,00	0	0	11	0	0	8	0	11	0	8	0
44	16/06/2017	6A	L	2	1	6	25 O	11,00	178	0	34	0	0	20	178	34	0	20	80
45	16/06/2017	6A	L	2	1	6	25 O	12,00	293	0	36	40	0	14	293	36	40	14	150
46	16/06/2017	6A	L	2	1	6	25 O	13,00	102	0	14	10	0	5	102	14	10	5	170
47	16/06/2017	6A	L	2	1	6	25 O	14,00	110	0	18	0	0	9	110	18	0	9	130
48	16/06/2017	6A	L	2	1	6	25 O	15,00	160	0	12	84	0	8	160	12	84	8	140
49	16/06/2017	6A	L	2	1	6	25 O	16,00	14	0	5	30	0	2	14	5	30	2	110
50	16/06/2017	6A	L	2	1	6	25 O	17,00	0	0	2	56	0	1	0	2	56	1	70
51	16/06/2017	6A	L	2	1	6	25 O	18,00	2	0	1	35	0	2	2	0	35	2	30
52	16/06/2017	6A	L	2	1	6	25 O	19,00	0	0	1	0	0	10	0	1	0	10	0
53	18/06/2017	6A	W	2	2	9	27 SO	10,00	110	16	44	0	0	27	126	44	0	27	75
54	18/06/2017	6A	W	2	2	9	27 SO	11,00	290	12	30	25	5	15	302	30	30	15	190
55	18/06/2017	6A	W	2	2	9	27 SO	12,00	198	21	43	48	8	15	219	43	56	15	220
56	18/06/2017	6A	W	2	2	9	27 SO	13,00	230	20	29	71	4	14	250	29	75	14	207
57	18/06/2017	6A	W	2	2	9	27 SO	14,00	61	20	29	191	27	9	81	29	218	9	197
58	18/06/2017	6A	W	2	2	9	27 SO	15,00	108	12	32	173	16	13	120	32	189	13	193
59	18/06/2017	6A	W	2	2	9	27 SO	16,00	56	17	19	225	13	14	73	19	238	14	173
60	18/06/2017	6A	W	2	2	9	27 SO	17,00	11	5	17	343	12	9	16	17	355	9	104
61	18/06/2017	6A	W	2	2	9	27 SO	18,00	8	0	7	244	0	2	8	7	244	2	14
62	18/06/2017	6A	W	2	2	9	27 SO	19,00	0	0	0	0	0	8	0	0	8	0	0
63	21/06/2017	6A	L	2	1	8	24 O	10,00	0	0	10	0	0	4	0	10	0	4	0
64	21/06/2017	6A	L	2	1	8	24 O	11,00	53	0	17	0	0	10	53	17	0	10	80
65	21/06/2017	6A	L	2	1	8	24 O	12,00	149	0	19	4	0	10	149	19	4	10	165
66	21/06/2017	6A	L	2	1	8	24 O	13,00	123	0	14	108	0	5	123	14	108	5	110

Figura 2. Matrice per la catalogazione delle variabili dipendenti ed indipendenti

Il primo output derivato dall'applicazione di RF riguarda l'importanza relativa di ogni variabile esplicativa per la determinazione dell'andamento della risposta osservata (Figura 3). Tutte le variabili osservate (presenze, arrivi e partenze) sono preliminarmente influenzate dalla fascia oraria che rappresenta la variabile nettamente più significativa in tutti i casi. Relativamente alle presenze, indipendentemente se in spiaggia o in baia, giocano un ruolo importante nel determinare il risultato anche il mese e la temperatura. La risposta è maggiormente eterogenea se si analizzano arrivi e partenze con l'eccezione delle condizioni del mare che tendono ad influenzare poco la risposta in tutti i casi analizzati. Analogamente anche l'intensità del vento è relativamente poco importante per gli arrivi e le partenze, ad eccezione degli arrivi via terra che invece sono significativamente influenzati da questo parametro.

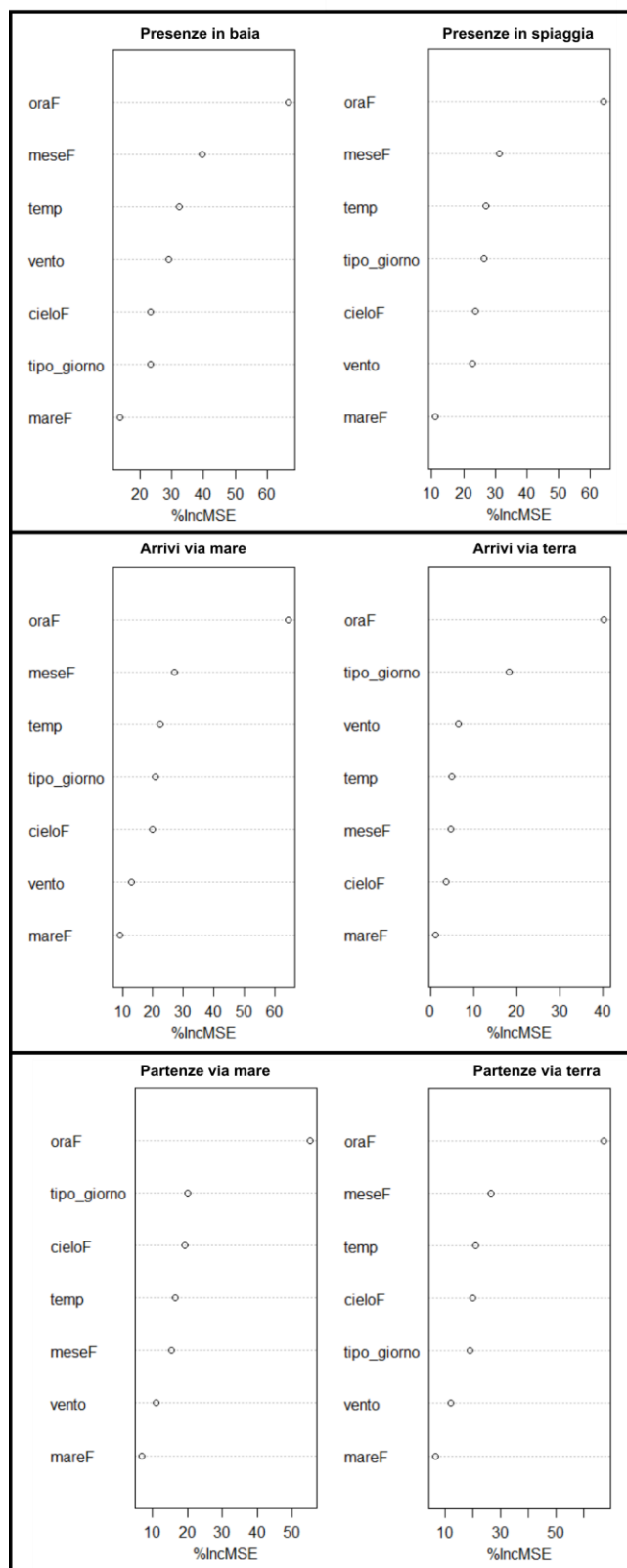


Figura 3: Diagrammi di importanza relativa delle variabili esplicative per ogni risposta ottenuta

Il tipo di dipendenza di ciascuna risposta dalle principali variabili è mostrato nelle Figura 4, Figura 5, Figura 6, Figura 7, Figura 8 e Figura 9. Il massimo di presenze in baia e in spiaggia è rilevato nelle ore centrali della giornata a seguito di arrivi concentrati in tarda mattinata e partenze che raggiungono il picco nella fascia oraria delle 16 sia via terra sia via mare. Il mese che registra le

presenze e gli spostamenti più elevati è agosto in concomitanza con le giornate del fine settimana in cui si rileva un generalizzato aumento dei segnali.

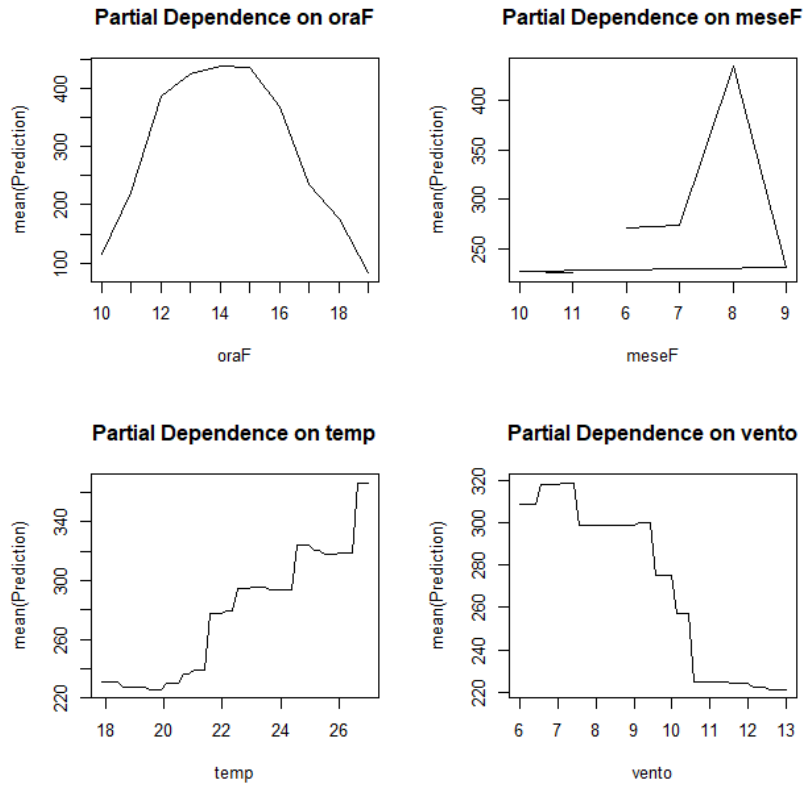


Figura 4: Dipendenza delle presenze in baia in relazione alle principali forzanti

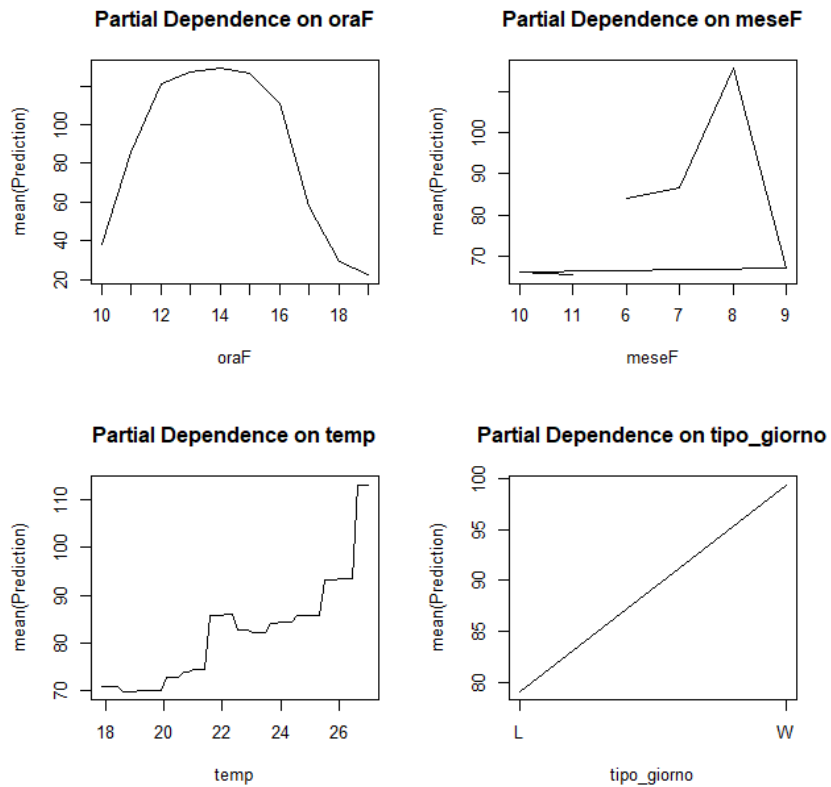


Figura 5: Dipendenza delle presenze in spiaggia in relazione alle principali forzanti

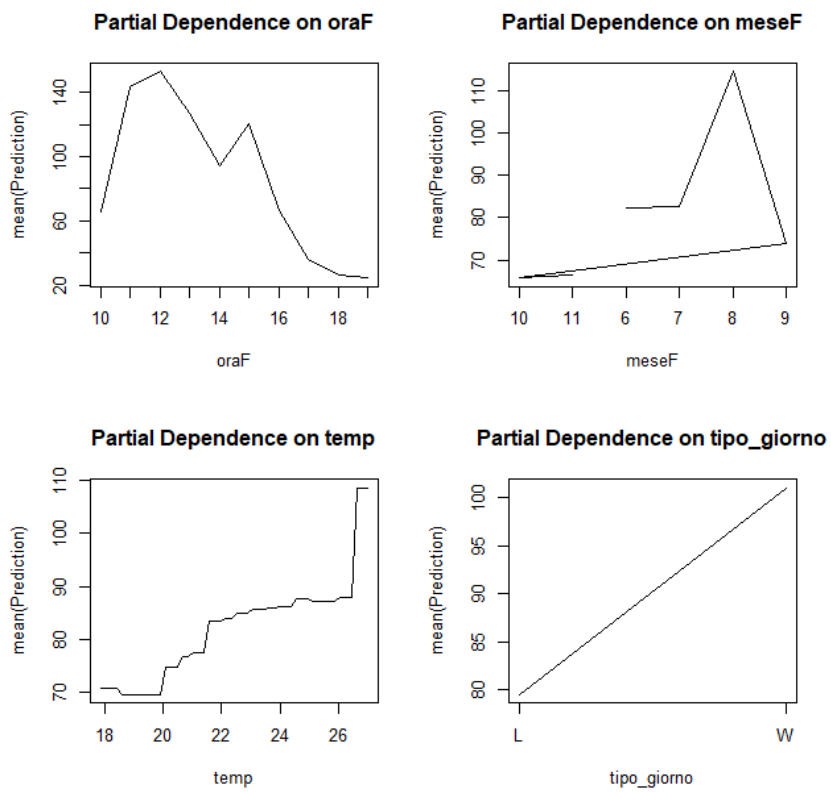


Figura 6: Dipendenza degli arrivi via mare in relazione alle principali forzanti

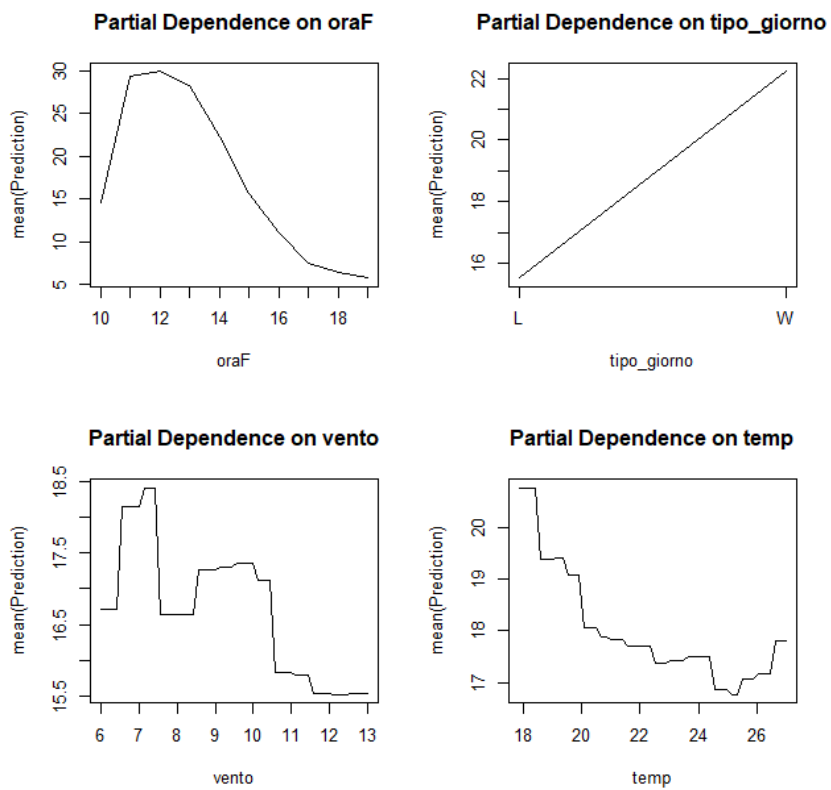


Figura 7: Dipendenza degli arrivi via terra in relazione alle principali forzanti

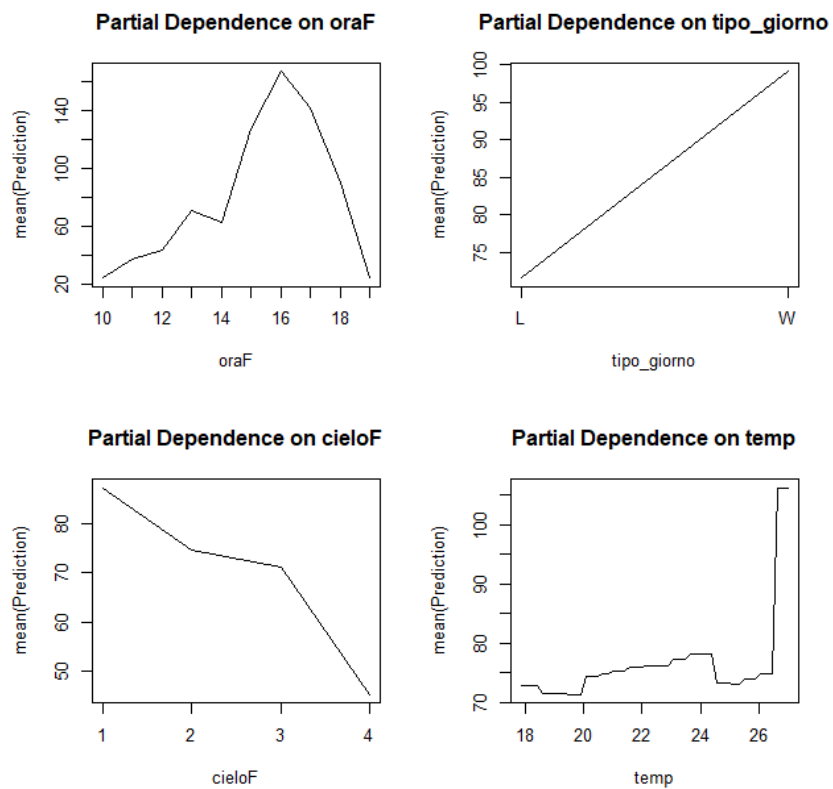


Figura 8: Dipendenza delle partenze via mare in relazione alle principali forzanti

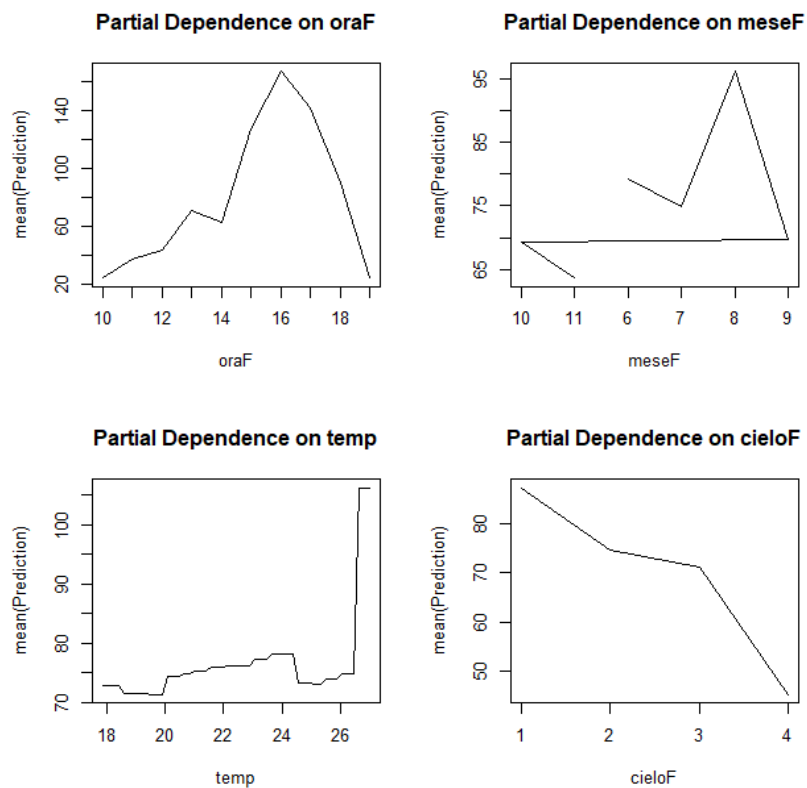


Figura 9: Dipendenza delle partenze via terra in relazione alle principali forzanti

Le regole di regressione ottenute dallo sviluppo del modello possono essere utilizzate al fine di predeterminare i flussi attesi a San Fruttuoso in funzione delle condizioni a contorno.

Per questo scopo è stato predisposto un foglio Excel (Figura 10) con una maschera di inserimento dei parametri necessari alla stima (mese, tipo di giorno, condizioni del cielo, condizioni del mare, intensità del vento, temperatura e ora). Tutti i parametri sono facilmente determinabili anche in condizioni di “forecasting” (previsioni meteo) e permettono una previsione del possibile affollamento e dei flussi a fini valutativi e gestionali.

Mese	tipo_giorno	cielo	mare	vento	temperatura	ora	Ora	Baia	Spiaggia	Arr Mare	Arr Terra	Par Mare	Par Terra	Arrivi	Partenze
							0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							10	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							11	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							12	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							13	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							14	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							15	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							16	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							17	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							18	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
							19	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A

Figura 10: Maschera di inserimento dei dati per la previsione dei valori di affollamento e dei movimenti su base oraria

2 Bibliografia

- Breiman L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Efron B. (1979). Computers and the theory of statistics: thinking the unthinkable. *SIAM review*, 21(4), 460-480.
- Hastie T., Tibshirani R., Botstein D., Brown P. (2001). Supervised harvesting of expression trees. *Genome Biology*, 2(1), research0003-1.
- Hastie T., Tibshirani R., Friedman J. (2009). Random forests. In *The elements of statistical learning* (pp. 587-604). Springer, New York, NY.
- Kehoe M., O'Brien K., Grinham A., Rissik D., Ahern K. S., Maxwell P. (2012). Random forest algorithm yields accurate quantitative prediction models of benthic light at intertidal sites affected by toxic *Lyngbya majuscula* blooms. *Harmful algae*, 19, 46-52.