

ALMA MATER STUDIORUM – Università degli Studi di Bologna

SCUOLA DI INGEGNERIA E ARCHITETTURA

**DIPARTIMENTO di INGEGNERIA CIVILE, CHIMICA, AMBIENTALE E DEI
MATERIALI - DICAM**

**CORSO DI LAUREA MAGISTRALE IN
INGEGNERIA PER L'AMBIENTE E IL TERRITORIO**

Tesi di Laurea Magistrale in:

Modellistica Idrologica

**Studio sperimentale del fattore di ragguglio delle
precipitazioni nell'Appennino Emiliano-Romagnolo**

CANDIDATA:

GIUDITTA SMERILLI

Relatore:

Chiar.mo Prof. Ing. Attilio Castellarin

Sessione II

Anno Accademico 2022-2023

Sommario

Introduzione	1
Capitolo 1: Approcci interpretativi e strumenti di calcolo per il fattore di riduzione delle piogge all'area	4
1.1 Fenomenologia e misurazione delle precipitazioni.....	4
1.2 Dati pluviometrici di riferimento	7
1.3 Il ragguaglio delle precipitazioni all'area	10
1.4 Approcci di valutazione del fattore di ragguaglio areale	13
1.5 Procedure di valutazione e formule di letteratura	16
1.5.1 La formula dell'U.S. Weather Bureau	16
1.5.2 La formula del N.E.R.C.	18
1.5.3 Le formule di Columbo e Puppini	21
1.5.4 Lo studio di Penta e la procedura di Battista, Pascuzzi e Catalano	22
1.5.5 La procedura di Moisello-Papiri	25
Capitolo 2: Dataset meteo orario e giornaliero regionale, Eraclito e ERG5	28
2.1 Banca dati Eraclito ed ERG5	28
2.2 La costruzione del dataset.....	30
2.3 La valutazione dei dati	31
2.3.1 I controlli di qualità.....	32
2.3.2 I controlli di omogeneità.....	36
2.3.2.1 Standard Normal Homogeneity Test	38
2.3.2.2 Test di Craddock	40
2.3.2.3 Test di Vincent	41
2.3.3 I controlli di sincronicità.....	42
2.3.4 Altri controlli	44
2.4 L'interpolazione	46
2.4.1 La griglia	46

2.4.2	Metodo di gridding e distanza topografica	48
2.5	Eraclito v. 3.0	51
2.5.1	Mappe climatiche	53
2.5.2	Valutazione degli errori	55
Capitolo 3: Tecnica sperimentale per la determinazione del fattore di ragguglio areale delle precipitazioni		59
3.1	Procedura sperimentale di valutazione dell'ARF	59
3.1.1	Dataset base	61
3.1.2	Dataset di interesse	63
3.1.3	Aggregazione delle altezze di precipitazione.....	64
3.1.4	Media dei massimi annuali	64
3.2	Il calcolo dell'ARF	65
3.3	Determinazione della pioggia media puntuale su un'area assegnata.....	66
3.4	Dati di precipitazione a passo orario ERG5	69
Capitolo 4: Risultati: il fattore di riduzione areale $ARF(A, d)$ sui bacini esaminati.....		70
4.1	Risultati sperimentali	70
4.1.1	Risultati sperimentali: $ARF(A, d)$ medio annuale	71
4.1.2	Boxplot della distribuzione dell' $ARF(A, d)$ medio annuale	74
4.2	Calibrazione di alcune leggi interpretative del fattore di riduzione areale $ARF(A, d)$ sui bacini esaminati.....	80
4.2.1	Metriche di prestazione considerate.....	85
4.2.1.1	R^2 modificato	85
4.2.1.2	Adjusted NSE.....	86
4.2.1.3	Bias relativo	88
4.1.1	Formula dell'USWB	88
4.1.2	Formula di Moisello e Papiri	95
4.1.3	Formula di Columbo	101
4.2	Conclusioni	108

Capitolo 5: Valutazione della dipendenza del coefficiente di ragguglio delle piogge all'area dal tempo di ritorno con riferimento agli eventi alluvionali del maggio 2023 in Emilia-Romagna.....	109
5.1 Gli eventi alluvionali del maggio 2023 in Emilia-Romagna	109
5.1.1 Gli eventi metereologici dell'1-4 maggio	110
5.1.2 Gli eventi metereologici dell'16-18 maggio	112
5.1.3 Le macro-aree regionali ARPAE	117
5.2 Confronto: ARF(A, d) annuale del 2023	118
5.2.1 Altezza di precipitazione media areale massima annuale	118
5.2.1.1 Plotting Position Correlation Coefficient.....	121
5.2.1.2 Test d'ipotesi.....	122
5.3 Boxplot.....	126
Conclusioni.....	131
Appendice.....	135
RStudio script per la valutazione dell'ARF annuale	135
RStudio script delle functions delle formule di letteratura calibrate.....	148
Andamento dell'ARF medio in funzione dell'area e della durata.....	153
Boxplot dell'ARF annuale	156
Bibliografia.....	167

Introduzione

La determinazione delle portate con un assegnato livello di rischio in una specifica sezione di un corso d'acqua, la progettazione di opere idrauliche, ecc., richiedono generalmente la valutazione preliminare della sollecitazione meteorica di progetto che interessa l'area oggetto di studio. La portata defluente in una sezione fluviale, inoltre, non dipende solo dalla durata e dall'entità dell'evento pluviometrico più critico verificatosi sul bacino sotteso, ma anche dalla posizione del centro di scroscio, dal suo spostamento e dalla distribuzione spazio-temporale della precipitazione. In particolare, la variabilità spaziale del campo di precipitazione, specialmente in corrispondenza di eventi di precipitazione significativi alla scala temporale subgiornaliera, è tanto più accentuata quanto maggiore è l'area in esame.

È quindi evidente che le valutazioni svolte facendo riferimento ai singoli pluviometri siano rappresentative di situazioni prettamente ed esclusivamente locali, ovvero ad una zona di influenza circoscritta al pluviometro considerato. Per utilizzare questo dato come dato di partenza, si potrebbe pensare di discretizzare il dominio in unità elementari corrispondenti alle aree di influenza delle singole stazioni, incrementando il livello di dettaglio e fornendo una rappresentazione differenziata del fenomeno di precipitazione sulla base dell'altitudine; d'altra parte, questi metodi sono caratterizzati da grande complessità e da tempi di calcolo più lunghi, per cui la loro applicazione è meno diffusa.

Di conseguenza, per avere un'informazione il più possibile rappresentativa della variabilità del fenomeno meteorologico spaziale e temporale sull'area considerata, si fa riferimento alla precipitazione media areale, determinata a partire da quella puntuale attraverso un coefficiente di riduzione o ragguglio areale, ARF (dall'inglese *Areal Reduction Factor*).

In letteratura, sono state proposte molteplici soluzioni per il ragguglio delle precipitazioni ad un'area specifica. È importante notare che, a causa del loro approccio spesso empirico, molte delle tecniche sviluppate sono strettamente legate al territorio considerato in sede di studio e, di conseguenza, possono essere applicate solamente in determinate regioni geografiche limitate.

Il presente studio si propone la stima sperimentale dell'ARF relativo ai massimi annuali di precipitazione per determinate zone e la calibrazione di alcune formulazioni di letteratura sui risultati ottenuti.

Nel dettaglio, la procedura adottata, tratta da studi storici sul fattore di riduzione della precipitazione all'area, fornisce la valutazione del fattore di riduzione all'area dei massimi

annuali di precipitazione puntuale ed areale annuale relative all'anno y per diverse durate dell'evento di pioggia d e per aree di riferimento coincidenti con i bacini idrografici ed i relativi sottobacini dei corsi d'acqua: F. Reno (ad esclusione degli affluenti romagnoli), F. Taro, F. Trebbia e F. Nure; con riferimento a un periodo di 32 anni, dal 1991 al 2022.

I dati pluviometrici di partenza impiegati sono stati derivati dal database ERG5, realizzato dal Servizio IdroMeteoClima di Arpae a partire dalle osservazioni raccolte dalla rete pluviometrica regionale. Il database consiste in un grigliato a maglia 5km di...

Ne conseguono la rappresentazione dei risultati ottenuti e la redazione di curve rappresentative del fattore di ragguglio areale della precipitazione in funzione dell'area e della durata considerata, trascurandone la dipendenza dal tempo di ritorno T .

Viene inoltre evidenziata la crescente variabilità del fattore di ragguglio areale con la precipitazione per ciascun bacino, considerandone la distribuzione nei 32 anni presi in esame.

Per ultimo, si esamina il risultato della valutazione dell'ARF annuale in funzione di area e durata, con riferimento al mese di maggio del 2023, periodo nel quale si sono verificati eventi alluvionali eccezionali per persistenza ed estensione spaziale nella porzione orientale della regione Emilia-Romagna, e a due delle macroaree regionali ARPAAE, particolarmente colpite da questi eventi. Tramite il confronto con i risultati precedenti e un approfondito studio del carattere eccezionale delle alluvioni con riguardo alla precipitazione media areale massima annuale ricavata per via sperimentale, si è cercato da un lato di quantificare l'effettiva uniformità spaziale delle piogge del maggio 2023, riscontrando valori di ARF particolarmente alti per diverse durate di riferimento e, dall'altro, di connotare la frequenza dei massimi di precipitazione areale durante l'evento al variare dell'intervallo di aggregazione temporale (durata).

L'elaborato è sviluppato come segue: dopo un breve excursus sulla fenomenologia delle precipitazioni, con particolare riferimento alla variabilità spaziale e temporale degli eventi di pioggia e al comportamento medio su larga scala e elevati orizzonti temporali, viene dedicata un'ampia sezione alla trattazione teorica dell'ARF, soffermandosi su studi storici a scala internazionale e italiana. Successivamente si descrive in maniera dettagliata il database pluviometrico di partenza, ERG5, dallo stato attuale al procedimento attuato dal SIMC ARPAAE per la sua realizzazione a partire dai dati storici della rete pluviometrica regionale. Viene quindi presentata la procedura di valutazione sperimentale applicata alle aree considerate, della quale vengono rappresentati e illustrati i risultati nel Capitolo 4. Dato l'evidente trend in comune ai valori del coefficiente di riduzione della precipitazione all'area

relativi a bacini appenninici differenti indagati, si è ritenuto opportuno calibrare alcune leggi empiriche, di importanza storica a livello mondiale e nazionale, sul campione sperimentale ottenuto. Sono descritte, nel medesimo capitolo, le performance dei vari modelli adattati ai valori empirici ricavati nel presente studio.

Infine, come caso di studio particolare, l'ultimo capitolo è dedicato alla connotazione delle precipitazioni medie areali del 2023, sia in termini di valori sperimentali di ARF, che di frequenza.

Capitolo 1: Approcci interpretativi e strumenti di calcolo per il fattore di riduzione delle piogge all'area

1.1 Fenomenologia e misurazione delle precipitazioni

La formazione delle precipitazioni è il risultato di meccanismi naturali cui consegue il raffreddamento della miscela aria-vapore acqueo, presente in atmosfera, che viene portata a saturazione. Questo fenomeno accade generalmente in contingenza alla migrazione di grosse masse di aria umida, che risalendo si portano a zone a bassa temperatura.

L'aria satura risalita si trasforma in nubi in corrispondenza di zone atmosferiche dette nuclei di condensazione, costituite da microparticelle di ossido di azoto, prodotti della combustione o sostanze saline, suscettibili di determinare la formazione di stato liquido. La fase di condensazione è seguita dalla fase di nucleazione nei nuclei di refrigerazione, in cui le goccioline d'acqua, precedentemente formatesi, rimangono sospese nell'atmosfera a causa della turbolenza. Questi elementi crescono in volume, principalmente attraverso il processo di coalescenza in cui un elemento più grande incorpora uno più piccolo, ma anche per il meccanismo di crescita per diffusione, dovuto ai gradienti di pressione di vapore presenti all'interno della nube. Quando le particelle diventano abbastanza grandi, le forze gravitazionali prevalgono e le fanno precipitare verso il basso.

Per loro natura le precipitazioni sono fenomeni intermittenti, enormemente mutevoli, la cui variabilità è funzione della tipologia in cui ricadono; queste possono infatti essere classificate in riferimento dinamica delle masse d'aria da cui hanno origine. In dettaglio, sulla base delle cause che hanno prodotto il sollevamento dell'aria, si distinguono le precipitazioni di tipo convettivo, ciclonico e orografico.

La risalita per convezione si verifica a causa di correnti convettive generate dai gradienti termici o di massa presenti localmente: un incremento della temperatura dell'aria in prossimità del suolo dà inizio a un movimento ascendente, convezione, durante il quale la miscela gassosa si espande e si raffredda progressivamente. Ne conseguono eventi di pioggia brevi e intensi, che interessano aree limitate con scale spaziali dell'ordine del km, come nel caso dei temporali estivi convettivi localizzati, quindi fenomeni variabili nello spazio che noi cogliamo, per quanto siamo capaci, attraverso delle reti di sensori.

Il sollevamento può avvenire anche per effetto di barriere orografiche, che ostacolano il passaggio dell'aria che si trova a transitare in loro prossimità, forzandola verso l'alto. La

miscela aria-vapore va incontro a zone a bassa pressione e bassa temperatura, riduzioni che favorisce la condensazione e la precipitazione; da questo fenomeno risultano infatti abbondanti eventi di pioggia.

Si può avere anche che la formazione di precipitazione sia favorita dallo scontro di masse d'aria con caratteristiche diverse a causa di un fenomeno di convergenza verso un'area a bassa pressione. Questo porta l'aria a muoversi verso l'alto (moti ascensionali): si creano i sistemi frontali, dovuti alla generazione di un gradiente di pressione all'interfaccia tra masse di aria calda e masse di aria fredda; in cui il fronte caldo tende in genere a scorrere sopra il preesistente cuscino di aria fredda. Le precipitazioni che ne derivano sono dette precipitazioni stratiformi, e sono piuttosto estese lungo i fronti di perturbazione e di modesta intensità, spesso a carattere nevoso. L'eventuale irruzione di un fronte freddo provoca un sollevamento repentino dell'aria calda che porta a piogge più circoscritte nello spazio ed intense, con fenomeni di tipo temporalesco.

La tecnologia attuale per la misurazione delle piogge spazia dal classico pluviometro, che fornisce un'osservazione in un punto ben preciso nello spazio, alle rilevazioni radar, che producono una periferica più completa di un evento meteorico e della sua distribuzione su larga scala.

Nel presente lavoro di tesi, i dati di precipitazione impiegati derivano dall'elaborazione di misurazioni raccolte dalla rete pluviometrica regionale. I pluviometri sono strumenti molto semplici, costituiti da un recipiente che raccoglie la pioggia captata tramite un imbuto posto al di sopra di esso. Per questo motivo, i valori raccolti da queste stazioni di misura sono considerati di valenza puntuale, poiché appunto svolgendo valutazioni a scala di bacino o con riferimento all'ambito urbano, essendo l'imbocco del pluviometro di diametro ristretto, indicativamente intorno ai 20 cm, questo è assimilabile a un punto. Al fine di standardizzare le misure e renderle comparabili, le rilevazioni di precipitazione da questa raccolte ne forniscono una misura in termini di altezza di pioggia: dal rapporto tra il volume dell'acqua raccolta per la superficie di intercettazione si determina l'altezza di precipitazione h , generalmente espressa in mm, che rappresenta l'altezza della lama d'acqua che coprirebbe, in un determinato intervallo di tempo, una superficie orizzontale, qualora tutta l'acqua caduta sulla superficie fosse trattenuta, così da formare uno strato di spessore uniforme. Estendere quindi un'osservazione considerata puntuale ad un'area comporta delle incertezze, che crescono all'aumentare dell'estensione dell'area per la quale si assume che un valore puntuale abbia valenza areale.

Per avere una rappresentazione adeguata della distribuzione degli eventi di pioggia nello spazio all'interno di un'area, inoltre, sarebbe necessaria una disposizione spaziale degli strumenti ad elevata densità. Nell'ambito dello studio di eventi di pioggia di particolare entità e breve durata, che ricadono su un'area tanto ristretta quanto maggiore è l'intensità stessa dell'evento, per ottenere una valutazione quantitativa dell'evento meteorico è necessaria una rete pluviometrica con una copertura più fitta. Se si desiderasse, invece, indagare il totale annuo o mensile in un'area scelta, è possibile considerare adeguata una densità di stazioni sicuramente inferiore.

In altre parole, la disposizione degli strumenti è funzione dell'informazione che si vuole ottenere a partire dai dati da essi registrati ed anche dalla variabilità spaziale della precipitazione; volendo ricavare una stima di una precipitazione media estesa all'area, allora è necessario l'impiego di un numero elevato di pluviometri, in modo che la distanza tra stazioni adiacenti sia il più ridotta possibile. Quest'ultima, per l'appunto, incide sulla precisione della stima della pioggia media areale: l'errore nella sua determinazione cresce in funzione dell'area e della larghezza delle maglie della rete.

Considerando che ogni strumento è rappresentativo di una zona di influenza che lo circonda, per ottenere una rete di pluviometri con peso all'incirca uniforme è necessario un incremento del numero di strumenti e della densità della loro distribuzione nelle aree dove i totali di pioggia sono più alti. Nelle zone desertiche una densità bassissima, non accettabile altrove, si può considerare soddisfacente. Anche il numero di strumenti influisce sulla precisione della stima dell'afflusso meteorico e tale precisione corrisponde a quella di una stima della pioggia areale, la cui importanza è fondamentale per ulteriori valutazioni, come ad esempio il calcolo e la previsione delle portate di piena.

Molto spesso accade che i costi di installazione degli apparecchi e la loro manutenzione ed accessibilità, ostacolano una rete ravvicinata dei pluviometri, e questo riduce notevolmente la possibilità di eseguire studi della distribuzione spaziale e temporale sufficientemente rappresentativi della realtà.

Le altezze di precipitazione registrate dai sensori di una rete pluviometrica possono però essere impiegati come punto di partenza per la determinazione della precipitazione media areale annuale che interessa una data area, ad esempio un bacino.

La valutazione della pioggia areale è, inoltre, condizionata dalla durata di precipitazione: nel dettaglio, questa diventa più accurata all'incrementare della durata dell'evento pluviometrico. Gli eventi meteorici che si protraggono per lungo tempo sono più uniformi nello spazio,

molto più di quelli brevi, che, d'altra parte, hanno un centro di scroscio molto ristretto. Per questo motivo, gli errori di calcolo della pioggia areale assumono grande importanza non tanto nello studio delle medie mensili o annuali, quanto nello studio delle piogge di forte intensità, generalmente associate a fenomeni temporaleschi di breve durata. Infine, l'errore di stima della precipitazione areale è influenzato anche della stagione considerata, e, in particolare, è maggiori in estate che in inverno, a causa del carattere isolato degli acquazzoni estivi, come si è osservato in numerosi studi.

1.2 Dati pluviometrici di riferimento

La Figura 1.1 fornisce un'idea della piovosità media annuale in Italia, rappresentata dalla linea verde a tratto continuo, che è appunto la media di lungo termine. Ogni istogramma è il quantitativo d'acqua piovuto in un anno, cumulato a scala nazionale.

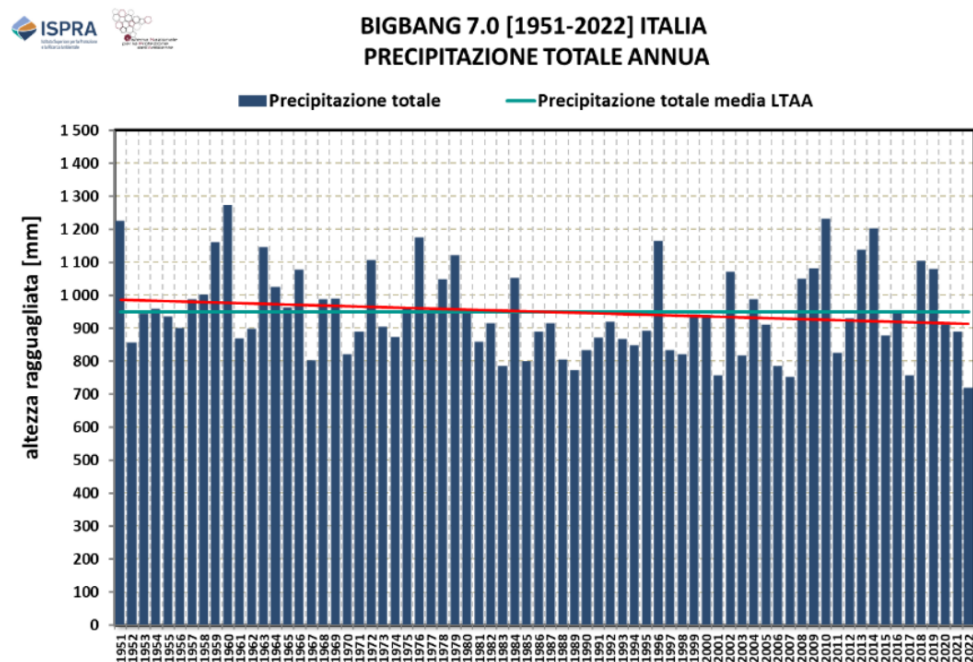


Figura 1.1 Serie storica delle precipitazioni annue dal 1951 al 2019 con l'indicazione della media LTAA e della media mobile (ISPRA).

Un valore medio della precipitazione in Italia di circa 1000 mm non fornisce però un'informazione utile per l'analisi delle precipitazioni, poiché queste ultime sono estremamente variabili nello spazio.

Osservando la Figura 1.2, in cui si evidenzia quanto accade ogni anno nello spazio in Italia in termini di precipitazione con riferimento alla media annuale di lungo termine:

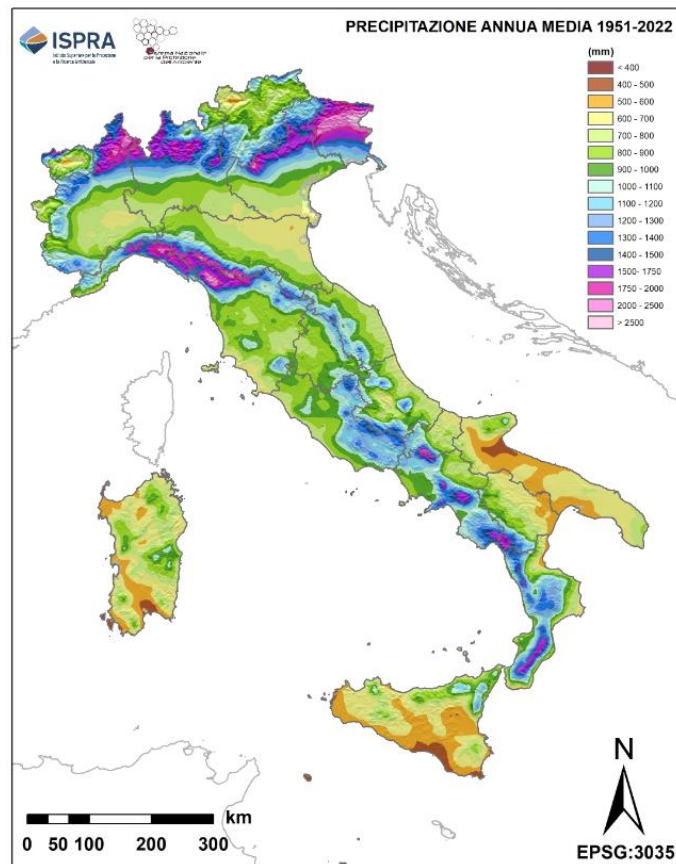


Figura 1.2 Precipitazione annua media, relativa al periodo 1951-2022 (ISPRA, su dati ufficiali dei servizi idrologici regionali e delle provincie autonome e del soppresso SIMN).

Si identificano zone in cui l'altezza di pioggia è ridotta (meno di 400 mm) che si trovano in porzioni meridionali del territorio, dovute a condizioni climatiche diverse, ovvero climi più caldi e secchi; d'altra parte, la pioggia media annuale in altre località può arrivare anche a superare i 2500 mm l'anno, più del doppio di ciò che cade a scala nazionale.

Osservando la mappa, si realizza immediatamente che sui totali annui di precipitazione hanno un impatto enorme le zone montuose: la componente orografica delle precipitazioni è legata alla presenza di catene montuose, che favoriscono la condensazione e la precipitazione. In Liguria, Toscana, si ha una piovosità prossima a quella che si osserva in Friuli: in corrispondenza di eventi estremi, si verificano piogge giornaliere, che scaricano al suolo 500 mm di precipitazione in 24 ore, metà di quanto piove a scala annuale su tutto il territorio italiano). Ciò è causato dalle correnti ricche di umidità provenienti dal Tirreno, che incontrano una prima barriera orografica, quella appenninica. Dopo essersi liberate in parte dell'umidità atmosferica, giunte quindi in Pianura Padana, sottovento, pioverà molto di meno; successivamente avviene l'incontro con una seconda barriera orografica, a quote superiori, che è l'arco alpino, quindi nuovamente si ha l'effetto orografico delle precipitazioni. Si osserva, inoltre, la formazione di una zona arida, con una forte variazione della precipitazione

media annua rispetto alla zona alpina vicina. Si tratta di una zona molto protetta e quindi, anche se alpina, questa è schermata da catene montuose dalle precipitazioni che arrivano da Sud e quelle che arrivano da Nord, e quindi vede sempre correnti di aria decisamente più secca, e si osserva una ridotta precipitazione rispetto al circondario.

Ragionare a scala annuale non permette di considerare che in realtà le precipitazioni sono estremamente variabili anche da un punto di vista stagionale, come illustrato in Figura 1.3.

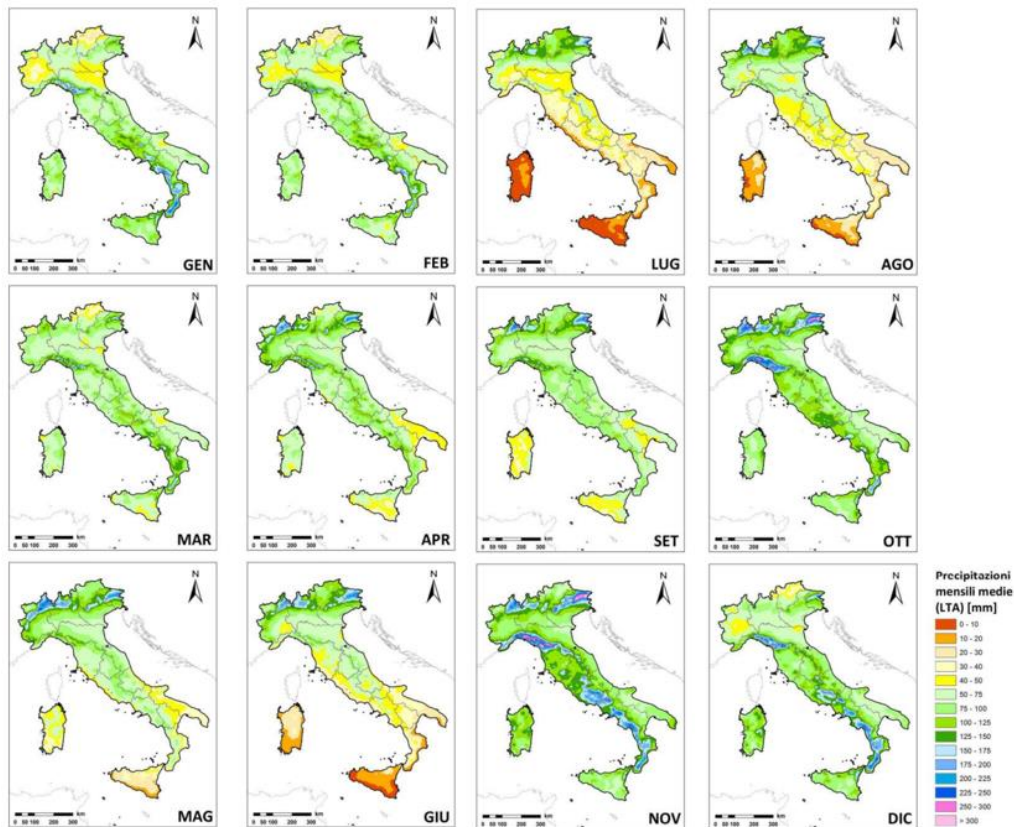


Figura 1.3 Precipitazioni mensili medie (LTA) (ISPRA).

Sebbene si osservi che i mesi meno piovosi sono quelli estivi, c'è una grandissima differenza tra Nord e Sud: nelle zone montane la piovosità non è ridotta come in quelle pianeggianti, e ci sono zone in cui è elevata (arco alpino) a causa di quei temporali estivi convettivi localizzati. I mesi più piovosi sono invece quelli autunnali, su tutta Italia e con una componente orografica marcata, tutto sempre ragionando in termini medi, con riferimento alle medie di lungo termine, ma la variabilità spaziale e temporale è elevata e, di pari passo, l'incertezza associata a valori medi di precipitazioni di valenza areale estesa. La variabilità spaziale infatti non è ben rappresentata da valori medi su un intervallo di tempo elevato.

1.3 Il ragguaglio delle precipitazioni all'area

Lo studio della variabilità spaziale e temporale dei fenomeni meteorologici, e in particolare delle precipitazioni, assume grande importanza per la comunità scientifica, specialmente in riferimento agli evidenti impatti economici e sociali causati da tali fenomeni sui sistemi umani.

La comprensione del comportamento a livello spaziale e temporale delle precipitazioni è fondamentale per la progettazione efficiente di infrastrutture idrauliche, non solo per uno sfruttamento ottimale delle risorse idriche per scopi civili, industriali ed agricoli, ma anche per proteggere le comunità dagli eventi meteorologici estremi e per mitigare i rischi idrogeologici. La struttura altamente variabile delle precipitazioni ha diverse implicazioni in molte questioni idrologiche, come la progettazione dei sistemi fognari, le misure di controllo delle alluvioni nei sistemi di drenaggio urbano, la progettazione di reti di sensori di pioggia e la previsione delle alluvioni, poiché è necessaria una risoluzione fine delle precipitazioni nello spazio e nel tempo. Il centro di scroscio, in dettaglio la sua posizione e il suo spostamento; e la distribuzione spazio-temporale della precipitazione influenzano fortemente le valutazioni che si svolgono a partire dai dati di pioggia.

Negli ultimi decenni, l'analisi delle precipitazioni è diventata una priorità, per via della crescente consapevolezza dell'innalzamento delle temperature e dei mutamenti climatici globali, con cambiamenti nella frequenza e nell'intensità degli eventi di precipitazione, cambiamenti che possono avere effetti significativi sulle risorse idriche, sull'agricoltura, sull'ambiente e sulla sicurezza delle comunità.

La variabilità spaziale, specialmente con riferimento ad eventi di pioggia particolarmente intensi, assume importanza all'aumentare dell'area in cui si verifica l'evento. È ampiamente riconosciuto che le precipitazioni si verificano, in linea generale, con significative variazioni nelle intensità su distanze inferiori a un chilometro: come dimostrato dalle osservazioni dei radar meteorologici, precipitazioni di diverse intensità, dimensioni e forme si verificano simultaneamente. Ciò è dovuto al fatto che fenomeni di precipitazione a vasta scala mesoscalare includono fenomeni a piccola scala mesoscalari, che a loro volta includono piccole celle di pioggia (Mineo et al.). Di conseguenza, solitamente le medie spaziali delle precipitazioni differiscono significativamente dalle misurazioni puntuali per vaste aree e brevi periodi.

Un'informazione il più possibile rappresentativa della variabilità del fenomeno sull'area considerata può essere ricavata dall'elaborazione opportuna di osservazioni contemporanee di

precipitazioni verificatesi per diversi eventi alle stazioni di misura. Essendo questa procedura troppo complessa, generalmente la pioggia areale viene stimata attraverso l'applicazione di un coefficiente di riduzione correttivo alla precipitazione puntuale. Il fattore che consente la trasformazione da precipitazione puntuale a precipitazione areale, è il coefficiente di riduzione areale (*Areal Reduction Factor*, ARF). L'ARF è definito come il rapporto tra le precipitazioni areali e le precipitazioni puntuali misurate da un pluviometro, ed è fondamentale per lo studio e la modellizzazione dei fenomeni idrologici.

Diversi elementi influenzano il rapporto tra la media spaziale delle precipitazioni su un'area specifica e le precipitazioni misurate in un punto, e possono essere correlati alle caratteristiche stesse delle precipitazioni, alle quelle del bacino in studio e ai dati e metodi utilizzati.

Il fattore di riduzione della precipitazione all'area è funzione della durata d , dell'area A e, debolmente, del tempo di ritorno T :

$$h_A(d, T) = h_p(d, T)ARF(d, A, T)$$

Dove $h_p(d, T)$ è l'altezza delle precipitazione puntuale, misurata da un pluviometro centrale durante un evento di pioggia intensa con durata d ; e $h_A(d, T)$ è l'altezza di pioggia media areale sull'area A alla durata d .

La stima del coefficiente di ragguglio areale è influenzata da:

- I processi di precipitazione: diversi autori hanno indagato la relazione tra il coefficiente di riduzione areale e i diversi processi di precipitazione. Ad esempio, secondo Skaugen (1997), il cui studio fa riferimento al Sud della Norvegia, l'ARF degli eventi di precipitazione su piccola scala spaziale registrati a passo giornaliero diminuiscono più rapidamente all'aumentare dell'area rispetto agli eventi di precipitazione frontali larga scala. Allen e DeGaetano (2005a) hanno riscontrato ARF più bassi in estate che in inverno: durante le stagioni calde, gli eventi di pioggia sono più convettivi e, quindi, caratterizzati da forti precipitazioni puntuali. D'altra parte, gli eventi di precipitazione estrema principali durante la stagione fredda sono essenzialmente sistemi frontali su larga scala con variabilità spaziale relativamente più piccola (rispetto alle tempeste convettive della stagione calda) che portano a un ARF più elevato. Diversi autori hanno segnalato una dipendenza tra l'ARF e la posizione geografica, che è correlata anche al clima e ai processi locali di precipitazione dominanti. La ricerca suggerisce anche una potenziale differenza nei ARF nelle aree urbanizzate

e in campagna, a causa della diversa densità della distribuzione della strumentazione per le aree rurali e urbane.

- Il tempo di ritorno: l'analisi della relazione tra l'ARF e il periodo di ritorno delle precipitazioni ha portato a risultati differenti. Alcuni autori hanno riportato che l'ARF diminuisce con l'aumentare del tempo di ritorno T , di solito a causa di un maggiore contributo dell'attività convettiva. Questi risultati sono in contrasto con altri studi, che non hanno trovato una relazione tra l'ARF e T per aree superiori a 500 km^2 . Ci sono state variazioni per aree più piccole, tuttavia, che gli Autori hanno spiegato con la bassa densità della rete di pluviometri, che non è in grado di rilevare eventi convettivi; e con la relativa breve durata dei registri di osservazione.
- Eventuali distorsioni nei dati delle precipitazioni utilizzati: il periodo della serie temporale delle precipitazioni utilizzato può influenzare gli ARF stimati a causa della variabilità temporale delle precipitazioni; inoltre, la combinazione di diverse reti di pluviometri per raggiungere una copertura spaziale accettabile può portare a un bias dei valori dell'ARF. La densità delle stazioni e le tecniche di interpolazione hanno tuttavia scarsa influenza sugli ARF, secondo Allen e DeGaetano (2005). Secondo i loro studi, inoltre, la stima del fattore di ragguaglio all'area può essere, in linea teorica, influenzata dall'altitudine: metodi di interpolazione come il metodo dei poligoni di Thiessen non tengono conto del fatto che i pluviometri sono più sparsi a diverse quote. D'altra parte, considerare questo effetto nelle tecniche di interpolazione spaziale non ha avuto un impatto significativo sugli ARF risultanti dalla loro analisi.
- Il metodo di stima; i possibili approcci per la stima del coefficiente di ragguaglio sono illustrati nei paragrafi successivi.

Inoltre, a rigore, il coefficiente di ragguaglio areale dipende dalla forma del bacino (Asquith, 1999); in genere, però, la dipendenza dalla forma si trascura.

La valutazione dell'ARF si realizza secondo numerosissime tecniche, sia metodologie valide su larga scala, cioè più generali, sia procedure applicabili a zone geomorfologiche ben definite, per le quali la stima è strettamente connessa al territorio considerato. In quest'ultimo caso, l'applicazione della metodologia e l'estensione dei risultati a zone diverse da quelle per cui il calcolo era stato concepito sono azzardate.

Sebbene questi metodi siano differenti tra loro, li accomunano caratteristiche generali, quali le

ipotesi di fondo: ci si aspetta che il valore riferito all'area sia inferiore, in termini di altezza di precipitazione, al valore puntuale; questo si traduce nell'assunzione seguente:

- Il pluviometro considerato si trova al centro dello scroscio dell'evento meteorico esaminato;
- L'evento pluviometrico è più intenso lì che altrove.

Da ciò si traggono due conclusioni:

- Il coefficiente ARF, $ARF = h_A / h$, assume valori tra 0 e 1;
- Ci si aspetta un andamento decrescente di ARF col crescere della superficie del bacino considerato.

Un'altra caratteristica comune è che a parità di area, il fattore ARF aumenta con la durata di precipitazione: più un evento è intenso, più è di breve durata e minore è l'estensione sul territorio, quindi è giustificata una riduzione maggiore del dato puntuale.

Dal punto di vista matematico, considerando le metodologie di calcolo dell'ARF e non quelle che forniscono in output una valutazione di una curva di possibilità climatica areale, è possibile schematizzare l'ARF come composizione di due funzioni, una dell'area sottesa e una della durata.

$$ARF = 1 - f_1(A)f_2(d)$$

Dove quindi:

- $f_1(A) = 0$ per $A = 0$;
- $f_1(A) = 1$ per $A \rightarrow \infty$;
- $f_2(d) = 1$ per $d = 0$;
- $f_2(d) = 0$ per $d \rightarrow \infty$.

Le espressioni dell' f_1 e dell' f_2 saranno diverse a seconda dell'autore che le ha stimate in funzione del bacino studiato.

1.4 Approcci di valutazione del fattore di ragguaglio areale

Poiché l'ARF è definito come il rapporto tra la media areale delle precipitazioni e le precipitazioni puntuali, esistono diversi approcci per la sua valutazione. Di conseguenza, diversi approcci possono risultare in ARF con diverse proprietà (Mineo et al.).

La stima del fattore di riduzione della precipitazione all'area può essere “ad area geograficamente fissa”, che descrive le caratteristiche della massima altezza delle precipitazioni in un bacino idrografico specifico (che può contenere solo una parte di un evento di pioggia o la somma di eventi multipli). La valutazione dell'ARF, d'altra parte, può essere “centrata sull'evento” che descrive l'evoluzione spazio-temporale di un singolo evento di pioggia per analisi deterministiche.

In generale, è risultato che il concetto di approccio ad area fissa è più coerente con l'analisi di frequenza delle piogge (*Precipitation Frequency Analysis*, PFA) e può fornire stime più conservative (Omolayo, 1993). Per questo motivo, l'ARF ad area fissa si usa generalmente nella progettazione delle reti di drenaggio, quello centrato sull'evento nella determinazione della precipitazione massima probabile (Barbero et al.). Il fattore di riduzione all'area, infatti, trova utilità non solo nell'ambito dei bacini naturali, ma anche in ambito urbano, per le reti artificiali, la cui progettazione deve essere riferita a un valore in ingresso opportunamente significativo.

Le prime ricerche si basano sull'analisi empirica di eventi di precipitazione singoli (*single storm events*), che hanno portato a formule che raramente tengono conto del periodo di ritorno dell'evento (U.S. Weather Bureau, 1958; Raudkivi, 1979). In Italia, ad esempio, tali formule si basano su alcuni studi pionieristici condotti nei primi decenni del secolo scorso (Supino, 1964) e, nonostante ciò, sono ancora molto diffuse nella definizione dell'evento pluviometrico di progetto per i sistemi di drenaggio urbano.

Un approccio diverso consiste in un metodo statistico per il calcolo dell'ARF di precipitazioni temporalesche estreme, basato sull'invarianza di scala temporale e spaziale e sull'autoaffinità statistica della funzione di distribuzione cumulativa (*Cumulative distribution Function*, CDF), dei massimi annuali delle precipitazioni puntuali, osservati presso la stazione di un pluviometro, e quella delle precipitazioni areali, ottenute mediando l'intensità delle precipitazioni su un'area circostante, A , per una determinata durata di pioggia, d . A condizione che il periodo di ritorno sia fisso, il fattore di riduzione è derivato come il rapporto dei quantili corrispondenti della CDF delle precipitazioni areali e quelli delle precipitazioni puntuali. Generalmente, questo tipo di metodo trascura l'influenza del periodo di ritorno (NERC, 1975; Moisello e Papiri, 1986); alcuni studi più recenti, tuttavia, mettono in evidenza una dipendenza minore, ma significativa, del fattore di riduzione dalla frequenza di comparsa dell'evento (Niemczynowicz, 1982; Stewart, 1989).

Negli ultimi decenni, è stata affrontata una diversa accezione del problema, seguendo un approccio stocastico; le funzioni di riduzione derivate si basano sulla struttura di correlazione spazio-temporale degli eventi pluviometrici. Purtroppo, come nelle analisi empiriche menzionate in precedenza, la funzione dell'ARF derivata in questo modo rappresenta la riduzione delle precipitazioni areali rispetto alle precipitazioni puntuali all'interno di un singolo evento meteorico. La stima necessaria a scopi di progettazione è invece la riduzione dell'intensità delle precipitazioni areali con frequenza assegnata, che può verificarsi in diversi eventi pluviometrici.

Veneziano e Langousis (2005) hanno adottato un approccio multifrattale per calcolare l'ARF, mettendo in evidenza le sue caratteristiche di scala, sotto l'assunzione che le precipitazioni siano scala-invarianti nello spazio e nel tempo e considerando anche l'influenza della forma del bacino considerato.

In conclusione, da quando è stato pubblicato lo studio empirico dell'U.S. Weather Bureau, sono stati proposti diversi metodi ARF, che possono essere classificati in:

- Metodi empirici: i metodi empirici hanno avuto origine dall'approccio dell'U.S. Weather Bureau e adattano una relazione statistica agli ARF sperimentali raccolti dall'area di studio. Nella pratica idrologica si fa spesso riferimento a studi empirici, che però sono fortemente dipendenti dalle condizioni geografiche e climatiche delle zone per le quali sono stati elaborati.
- Metodi della Teoria dei Valori Estremi: estendono la distribuzione dei valori estremi, ampiamente utilizzata nell'analisi delle frequenze dei valori puntuali, alle precipitazioni estreme areali.
- Metodi di correlazione spaziale: si basano sulla correlazione spaziale dei campi di precipitazione, sotto l'ipotesi che la correlazione spaziale non varia significativamente lungo una direzione specifica, e particolari distribuzioni statistiche del processo di precipitazione.
- Metodi di scaling spaziale e temporale: utilizzano i concetti di scaling dinamico, multifrattale e autosomiglianza statistica per riflettere le proprietà di scaling delle precipitazioni nello spazio e nel tempo.

Di seguito, vengono analizzate le formule utilizzate in sede di calibrazione dei risultati ottenuti e gli approcci di valutazione dell'ARF su cui si basa la procedura sperimentale applicata.

1.5 Procedure di valutazione e formule di letteratura

1.5.1 La formula dell'U.S. Weather Bureau

Tra il 1957 e il 1960, l'U.S. Weather Bureau (USWB), il Servizio Meteorologico degli Stati Uniti, ha raccolto ed elaborato dati di precipitazione al fine di fornire una relazione empirica per la stima del fattore di riduzione delle piogge all'area.

Il Technical Paper No. 29 (TP-29, U.S. Weather Bureau, 1958) rappresenta il primo sviluppo sistematico degli ARF per gli Stati Uniti, fornendo stime per durate che vanno da 30 minuti a 24 ore e aree dai 250 fino a 1036 km². L'analisi del TP-29 si basava sulle osservazioni di vari eventi piovosi registrati in sette stazioni provenienti da fitti network di pluviometri dispiegati negli Stati Uniti; per un periodo variabile dai 7 ai 15 anni circa. Principalmente a causa della limitata durata dei dati e delle piccole aree di studio, le curve ARF del TP-29 non caratterizzano alcuna dipendenza dal periodo di ritorno, dalla stagione o dalla posizione geografica. Anche se il TP-29 è stato pubblicato decenni fa, è ancora utilizzato in molte applicazioni ingegneristiche oggi, anche se le curve sviluppate possono differire significativamente quando valutate in diverse parti del mondo.

Sostanzialmente, lo studio si sviluppa come segue: per una data area, dalle serie storiche a disposizione si estraggono i dati di precipitazione massima puntuale per ogni durata considerata rilevati dalle stazioni. Nel caso in cui, auspicabilmente, siano presenti un numero di pluviometri maggiore di uno nell'area in esame, si ricava la media pesata dei massimi puntuali, utilizzando, ad esempio, il metodo dei topoietai o delle isoiete per la valutazione dei pesi.

Si passa quindi al calcolo delle precipitazioni medie areali, a partire dalle serie puntuali, come la media pesata tramite l'uso di opportuni pesi che permettano di ricavare un valore di riferimento "distribuito". I pesi, anche in questo caso, possono essere stimati tramite metodi come il metodo dei poligoni di Thiessen. Si procede quindi con l'aggregazione secondo le varie durate.

A partire dalle serie storiche dei dati medi puntuali e dei dati medi areali per ogni durata considerata, dal confronto dei valori massimi di assegnato tempo di ritorno estratti da queste, si definisce il fattore di riduzione di ragguaglio all'area:

$$ARF(A, d, T) = \frac{h_A(d, T)}{h_P(d, T)}$$

Dove:

- $h_P(d, T)$ rappresenta il valore medio massimo puntuale per assegnati durata d e tempo di ritorno T ;
- $h_A(d, T)$ corrisponde al massimo areale, per uguali durata e tempo di ritorno.

Nel 1972 Eagleson ha formulato una relazione dell'ARF in funzione della durata d della precipitazione e dell'area su cui si vuole spazializzare la misura puntuale, previa considerazione della dipendenza dal tempo di ritorno T come trascurabile. Secondo questa, l'andamento di ARF in funzione di d ed A ha espressione analoga alla:

$$ARF = 1 - f_1(A)f_2(d)$$

In dettaglio, f_1 ed f_2 , secondo Eagleson, hanno le seguenti forme:

$$f_1(A) = 1 - \exp(-c_1 A)$$

$$f_2(d) = \exp(-c_2 d^{c_3})$$

In conclusione, gli studi empirici dell'USWB hanno portato come risultato alla seguente relazione dell'ARF:

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

Dove d e A sono la durata e l'area e c_1 , c_2 e c_3 sono i tre parametri. Questa relazione rispetta tutte le condizioni precedentemente illustrate, per qualunque valore di A e di d .

Di seguito, sono riportate le curve proposte dall'USWB (Eagleson) per la determinazione dell'ARF (Figura 1.4). È importante tenere presente che queste curve sono state calibrate per il territorio degli USA, e quindi vanno impiegate con cautela in applicazioni su quello italiano, in quanto caratterizzato da condizioni meteo-climatiche distinte.

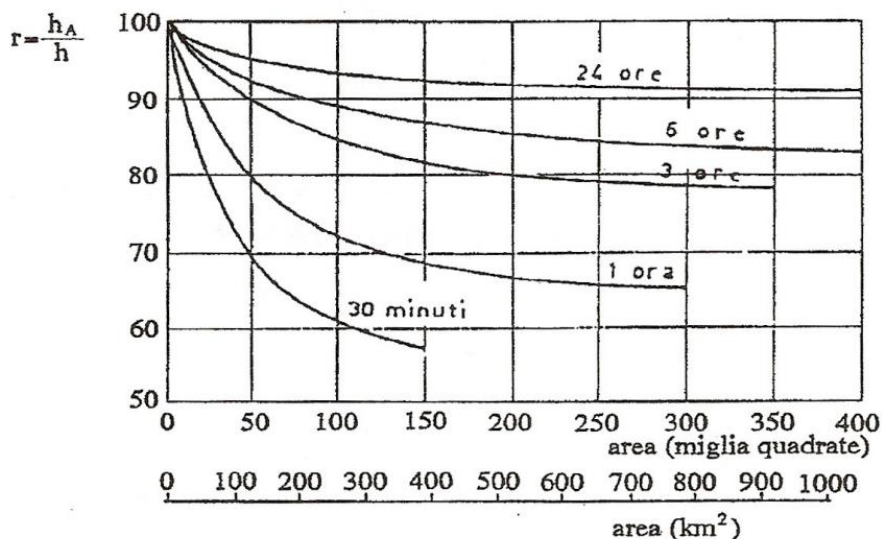


Figura 1.4 Curve ricavate da Eagleson, per il territorio statunitense.

Gli studi empirici dell'U.S. Weather Bureau sono stati seguiti dallo sviluppo di nuove formulazioni da parte di diversi autori. Di seguito si riporta la formulazione empirica proposta dal Natural Environmental Research Council (N.E.R.C.), per via dell'importanza storica, e le espressioni calibrate in sede di analisi sui bacini appenninici tosco-emiliani.

1.5.2 La formula del N.E.R.C.

Il metodo ad area fissa attualmente utilizzato nel Regno Unito, presentato nel Flood Studies Report (NERC 1975), non tiene conto del periodo di ritorno, poiché se ne assume che l'effetto fosse trascurabile.

Il metodo di valutazione consiste nella stima, per un'area A e per una durata d , della massima piovosità areale annuale. In altre parole, viene identificato l'evento areale più gravoso, responsabile della maggiore altezza di precipitazione media areale, h_A , massima. In corrispondenza di tale evento areale, si annotano, per ogni punto dell'area in cui è collocato uno strumento, le misurazioni puntuali, $h_P'_{ij}$, presso la stazione i e nell'anno j .

Si procede poi prendendo le massime piogge puntuali annuali, $h_{P,ij}$, presso ciascuna stazione i per l'anno j , nella regione. Per ciascuna area di superficie A e per ciascuna durata d , l'ARF viene calcolato come la media dei rapporti tra la pioggia caduta durante l'evento in un punto e il massimo annuale nello stesso punto, eseguita spazialmente su tutti i punti dell'area e temporalmente su tutti gli anni di osservazione:

$$ARF = \frac{1}{IJ} \sum_j \sum_i \frac{h'_{p,ij}}{h_{p,ij}}$$

Dove I è il numero totale di stazioni nella regione e J è la lunghezza del record (anni). Questa relazione è stata impiegata per il calcolo del fattore di ragguglio areale per tredici bacini nel Regno Unito, con aree che vanno da 10 a 18000 km², per durate medie variano da 2 minuti a 25 giorni. Successivamente, il N.E.R.C. ha elaborato e interpolato i risultati empirici per ottenere il diagramma nella Figura 1.5.

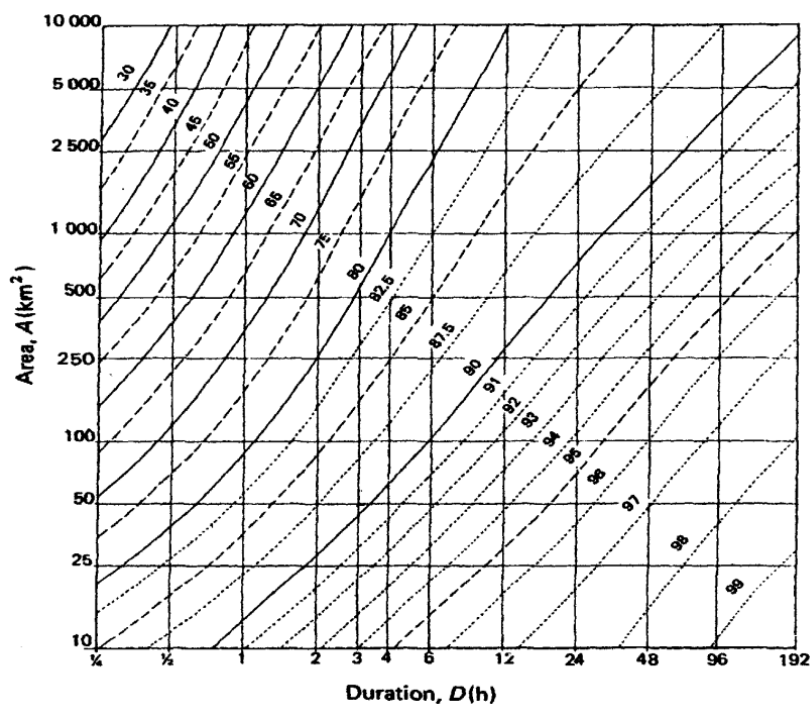


Figura 1.5 Fattori di riduzione areali (ARF) per durate D dai 15 minuti agli 8 giorni e aree A dai 10 ai 10000 km². Riprodotto da N.E.R.C. (1975) Flood Studies Report, Vol. II, p. 40.

Dall'interpolazione dei risultati sperimentali ottenuti dal N.E.R.C., si osserva che l'ARF è approssimativamente costante per durate proporzionali alla potenza dell'area, $d \propto A^{0.7}$; tuttavia, le curve iso-ARF del N.E.R.C. indicano una pendenza variabile, con $d \propto A^{0.5}$ per A grandi e d ridotte, e esponenti A maggiori per aree più basse o d più grandi.

In sostanza, il metodo proposto da N.E.R.C. è una semplificazione del metodo del Weather Bureau degli Stati Uniti descritto in precedenza, che si basa sull'assunzione che “una media di rapporti” possa approssimare in modo sicuro “un rapporto di medie” (N.E.R.C., 1977), il che può essere considerato piuttosto non convenzionale.

Nel 1976, Bell ha eseguito la revisione e la rivalutazione dei risultati ottenuti nel Flood Studies Report, considerando, tra le altre cose, all'assunzione che gli ARF siano indipendenti

dal periodo di ritorno. Lo ha fatto considerando l'applicazione delle distribuzioni di frequenza alle serie di massime precipitazioni annuali areali e puntuali, e quindi calcolando l'ARF come rapporto tra le stime di altezza di pioggia areale e puntuali dello stesso tempo di ritorno. Il metodo prevede il posizionamento in ordine crescente delle precipitazioni medie areali massime annuali (ponderate secondo il metodo dei poligoni di Thiessen), e quindi anche il posizionamento delle precipitazioni puntuali massime annuali, presso ciascuna stazione presente nell'area. Per ottenere una singola curva di frequenza delle piogge puntuali rappresentativa per l'area, è stata calcolata la media ponderata delle massime piovosità puntuali annuali dello stesso rango, secondo il metodo dei topoi. Sono state quindi adattate distribuzioni di frequenza alle serie areali e puntuali così ottenute e calcolati gli ARF per diversi periodi di ritorno. Bell (1976) ha individuato casi in cui l'ARF che diminuisce più rapidamente con l'aumentare del periodo di ritorno, ma ha concluso che l'uso degli ARF calcolati secondo il Flood Studies Report (NERC 1975) risultava in stime conservative degli ARF.

Bell si occupò anche del confronto dei risultati ottenuti il metodo di N.E.R.C. con quelli ricavati tramite la procedura sviluppata dell'USWB; osservando che solo per determinate tipologie di dati in ingresso i risultati coincidevano. Le stime prodotte dai due metodi sono concordi quando si considerano valori di precipitazioni medi e non corrispondenti a tempi di ritorno elevati, per i quali invece si evidenziava con una certa sistematicità che l'ARF ricavato tramite il metodo di N.E.R.C. era una sovrastima di quanto risultante dal metodo dell'USWB.

In seguito alla revisione di Bell, N.E.R.C. modificò i propri risultati, fornendo le conclusioni in forma grafica. Sulla base delle curve del N.E.R.C., l'Institute of Hydrology di Wallingford propone una formula del tipo:

$$ARF = 1 - f_3 d^{f_4}$$

Dove f_3 e f_4 sono funzioni dell'area A. Un'adeguata interpolazione delle valutazioni si ottiene anche con la relazione:

$$ARF = (1 + \bar{\omega} A^a d^{-b})^{-c}$$

Dove i parametri $\bar{\omega}$, a, b, c sono stimati a partire dalle curve prodotte dal N.E.R.C. e riportate nel Flood Studies Report, tali per cui:

$$ARF = (1 + 0,02A^{0,37} d^{-0,48})^{-2}$$

Con A in km² e d in ore.

1.5.3 Le formule di Columbo e Puppini

Lo studio del coefficiente di ragguglio areale nel contesto nazionale segue un approccio differente dai metodi finora trattati. Lo scopo dell'analisi non è quello di fornire una stima del coefficiente di riduzione della precipitazione puntuale all'area e quindi di stimarne la dipendenza dall'area e dalla durata. In Italia, le elaborazioni sono impostate al fine di ricavare un'espressione di una curva di possibilità climatica areale, modificando i coefficienti della curva di possibilità climatica puntuale, così da ricavare una formulazione della curva areale analoga a quella della curva puntuale, del tipo:

$$h_A(d, T) = a' d^{n'}$$

Con a' e n' coefficienti. Gli studi di Columbo e Pappi hanno portato a differenti espressioni di a' ed n' , che però sono accumulate dalla dipendenza dai coefficienti della curva climatica puntuale e dall'area A considerata.

Le formule di Puppini (1932), largamente utilizzate per il caso di piccoli bacini, sono state ricavate dall'autore per la Bonifica Renana, a partire da dati di precipitazione forniti da stazione ubicate a Milano e pubblicate da Massari nel 1910, valide per aree inferiori a 600 km² e per durate comprese tra 1 e 15 giorni:

$$a' = a \left[1 - 0,084 \left(\frac{A}{100} \right) + 0,007 \left(\frac{A}{100} \right)^2 \right]$$

$$n' = n + 0,014 \left(\frac{A}{100} \right)$$

Dove l'area A è riportata in km².

Le formule di Columbo (1960) state sviluppate nell'ambito dello studio delle precipitazioni intense nella città di Milano, per aree inferiori a 5000 ha e durate, in ore, inferiori a 24 h:

$$a' = a \left[1 - 0,06 \left(\frac{A}{100} \right)^{0,4} \right]$$

$$n' = n + 0,003 \left(\frac{A}{100} \right)^{0,6}$$

In cui l'area è espressa in ha.

Condizione per l'applicazione di entrambe le curve è che la curva di possibilità climatica sia ottenuta elaborando le osservazioni misurate in corrispondenza del centro di scroscio; inoltre, si presuppone, in entrambi gli studi, che il fattore di riduzione della precipitazione all'area

dipenda solo dall'area e dalla durata della precipitazione e non dalla sua intensità. Nel caso, non raro, in cui la curva puntuale sia redatta a partire dalle rilevazioni di un unico strumento di misura, l'ipotesi che il centro di scroscio coincida con la posizione dello strumento è arbitraria, così come l'assunzione che dell'indipendenza dal tempo di ritorno, ovvero dall'entità dell'altezza di precipitazione.

In entrambe le analisi, dal confronto dei coefficienti riferiti alle curve puntuale e areale, si ottiene:

- $a' < a$;
- $n' > n$.

D'altra parte, il coefficiente a ha il significato dell'altezza di precipitazione alla durata unitaria: il valore mediato sull'area è necessariamente inferiore a quello a valenza puntuale. Il coefficiente n è il coefficiente angolare della retta di possibilità climatica in scala semilogaritmica e ha il significato del tasso di crescita della massima altezza di pioggia in funzione della durata. Di conseguenza, dal momento che all'aumentare della durata il fattore di ragguaglio areale aumenta anch'esso e tende a 1, è naturale che la retta climatica areale abbia maggiore pendenza di quella puntuale.

1.5.4 Lo studio di Penta e la procedura di Battista, Pascuzzi e Catalano

A seguito della pubblicazione del Technical Paper No. 29 dell'U.S. Weather Bureau, nel 1974 Penta ha condotto uno studio sul fattore di ragguaglio areale per i bacini del versante ionico della Basilicata, considerando piogge giornaliere e ricavando una espressione dell'ARF in funzione dell'area. La scelta di limitare la durata giornaliera è dovuta al fatto che le serie storiche impiegate nell'analisi, venivano estratte dagli Annali Idrologici, nei quali sono raccolti i dati di pioggia giornaliera (intesa come l'altezza di precipitazione, in mm, caduta nelle 24 ore fra le 09:00 del giorno in cui si effettuava la misura e le 09:00 del giorno precedente).

La procedura elaborata da Penta, sebbene simile a quella dell'USWB, ha portato a valutazioni molto differenti sia dallo studio statunitense che da quello del N.E.R.C.; questo risultato può essere spiegato considerando che il fattore di riduzione areale è correlato alle caratteristiche meteorologiche dalla zona su cui viene valutato, e per questo motivo quanto ottenuto da Penta non è direttamente confrontabile con gli output di questi studi.

Questa incongruenza non si incontrerebbe se si considerasse di impiegare, piuttosto che un

metodo empirico, un approccio di tipo stocastico tenga conto della distribuzione spaziale della precipitazione attraverso la rappresentazione della struttura di correlazione spaziale alle diverse durate.

A partire dalle osservazioni di Penta, è stata formulata la seguente espressione dell'ARF:

$$ARF(A, d) = 1 - f_1(A)f_2(d)$$

Con:

$$f_1(A) = 1 - \exp(-c_1 A)$$

$$f_2(d) = \exp(-c_2 d^{c_3})$$

La relazione finale è:

$$ARF(A, d) = [1 - \exp(-0,0021A)][\exp(-0,53d^{0,25})]$$

Con A in km² e d in ore.

In Calabria, un'indagine sull'andamento areale delle precipitazioni giornaliere in riferimento al Fiume Crati, svolta da Battista (1988), Pascuzzi e Catalano (1988), ha fornito risultati sostanzialmente analoghi a quelli ottenuti per la Basilicata. Di particolare importanza è la seguente trattazione, in quanto questa metodologia è stata presa a riferimento per la stima dell'ARF per le aree prese in esame in sede di tesi, applicata per durate diverse da quella giornaliera.

I dati di riferimento sono altezze di precipitazione giornaliera provenienti da 34 stazioni del Servizio Idrografico e Mareografico Nazionale (SIMN), relativi ad un periodo che va dal 1925 al 1968. Il bacino idrografico del Fiume Crati ha un'estensione di 2356 km²; nell'analisi, sono stati considerati 8 sottobacini sottese da altrettante sezioni fluviali consecutive verso valle. Per ciascuna delle aree prese in esame è stata ricavata la serie storica di precipitazione media areale, impiegando il metodo dei poligoni di Thiessen, la cui attendibilità per reti pluviometriche a bassa densità è comparabile a quella di metodi più complessi.

Le variabili di interesse nella seguente metodologia sono:

- $h_{d,A}$, l'altezza di precipitazione giornaliera media areale sull'area di estensione A;
- $h_{d,i}$, l'altezza di precipitazione giornaliera puntuale registrata dall'i-esimo pluviometro sull'area A;
- n_A , il numero totale di pluviometri sull'area A;

- P_i , il peso dell' i -esimo pluviometro secondo il metodo dei topoi.

La precipitazione media areale giornaliera corrisponde alla media pesata dei valori puntuali dei singoli pluviometri, il cui peso sulla stima è funzione dell'area di loro competenza:

$$h_{d,A} = \sum_{i=1}^{n_A} P_i h_{d,i}$$

Per ogni sottobacino sono state ricostruite le serie storiche della precipitazione $h_{d,A}$, da cui sono stati estratti i massimi annuali delle precipitazioni medie areali. Dal calcolo della media temporale su tutti gli anni di osservazione, si ottengono i valori della media dei massimi della precipitazione areale $h_{d,A,m}$ per ciascun sottobacino di area A .

Le serie delle precipitazioni puntuali originali sono state nuovamente elaborate: sono stati estratti i massimi annuali puntuali per ogni pluviometro ricadente nell'area; dai massimi annuali, in numero pari al numero dei pluviometri per gli anni di osservazione considerati, è stata ricavata la loro media aritmetica, $\bar{h}_{d,i,max}$. Infine, ne è stata ottenuta la loro media ponderale, con riferimento agli stessi pesi impiegati per la precipitazione media areale:

$$h_{d,A}^* = \sum_{i=1}^{n_A} P_i \bar{h}_{d,i,max}$$

Il fattore di riduzione della precipitazione all'area è definito come:

$$ARF = \frac{h_{d,A}}{h_{d,A}^*}$$

Con riferimento alla formula di Eagleson:

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

Si può scrivere, considerando che c_3 è 0,25:

$$ARF = 1 - \exp(-c_2 d^{0,25}) + \exp(-c_1 A) \exp(-c_2 d^{0,25})$$

Inoltre, i dati di precipitazione sono presi solo per la durata giornaliera, per cui si può porre d costante e pari a 1, da cui:

$$ARF = a + (1 - a) \exp(-bA)$$

Dove:

$$a = 1 - \exp(-c_2)$$

$$b = c_1$$

Sostituendo con i parametri calibrati dagli Autori sul Fiume Crati con il metodo dei minimi quadrati, la scrittura definitiva per l'ARF è:

$$ARF = 0,727 + 0,273 \exp(-0,00169A)$$

Valida per una durata di 1 giorno.

La concordanza con i risultati dello studio di Penta sicuramente fornisce una validità più generale alla formulazione sopra descritta.

È evidente che le formule presentate abbiano forti limitazioni; ad esempio, non è possibile applicarle con precisione a regioni geografiche che hanno regimi pluviometrici differenti da quelli per i quali sono state tarate. Il limite più rilevante riguarda la mancanza di generalità delle formule stesse, in quanto si basano su leggi puramente empiriche, sprovviste di qualsiasi fondamento teorico.

1.5.5 La procedura di Moisello-Papiri

Un'altra, più recente rispetto alle altre, formulazione empirica del fattore di riduzione è stata proposta da Moisello & Papiri (1986) in base allo studio di un'area di circa 1000 km² centrata sulla città di Milano, caratterizzata da una rete di 16 pluviografi sufficientemente omogenea e da un regime pluviometrico non del tutto uniforme nello spazio. I dati di pioggia considerati fanno riferimento a un periodo di osservazione dal 1977 al 1981; ne risulta una formulazione valida per durate da 15 minuti a 12 ore e aree da 5 a 800 km².

Lo studio nasce dalla necessità di definire per la zona della città di Milano, un valore in ingresso per la progettazione di una rete di drenaggio, ad esempio un sistema di fognature urbane. Solitamente, la portata di ciascun tronco della rete si calcola facendo riferimento a un unico valore di precipitazione, assunto costante nel tempo e nello spazio. Questa scelta è legata sia alla complessità dei modelli a più ingressi, che alla difficoltà nel fornire una rappresentazione della variabilità spazio-temporale della precipitazione in termini probabilistici; sia per la carenza di dati pluviometrici necessari per ottenere una schematizzazione complessa. L'input deve essere, quindi, riferito all'intera area in esame e quindi deve essere quello medio areale.

I metodi empirici, come ad esempio quello di Columbo e Puppini, si basano sull'assunzione che la curva di possibilità climatica puntuale da cui si ottengono i coefficienti a ed n per la

valutazione della curva climatica areale, sia ricavata dall'elaborazione di dati di precipitazione raccolti in corrispondenza del centro di scroscio. Spesso però accade che la curva puntuale sia costruita a partire dai dati registrati da un solo pluviometro, e pertanto l'assunzione che la sua posizione coincida con il centro di scroscio è debole.

Moisello e Papiri hanno sviluppato una procedura statistica che si basa sull'introduzione del concetto di pluviografo centrale. Questa consiste nelle seguenti operazioni: nell'area assegnata si indica considera una stazione come centrale e si individua una zona circolare di raggio prestabilito avente come centro lo strumento scelto. Per ogni evento di precipitazione, di entità qualsiasi, viene stimato lo ietogramma della pioggia media areale, impiegando il metodo dei poligoni di Thiessen per la stima dei pesi. Si fa poi riferimento alla precipitazione puntuale, e, in particolare, si ricercano tra gli ietogrammi puntuali in corrispondenza del pluviografo centrale, gli intervalli temporali di ampiezza assegnata, assunta di volta in volta pari ad una delle durate considerate, ai quali corrispondono, rispettivamente, la precipitazione massima annuale puntuale e la precipitazione massima annuale areale. Tuttavia, al fine di ottenere un numero superiore di dati selezionati, oltre che a prendere i massimi annuali, sono evidenziano i diversi eventi di precipitazione rilevanti di ciascun anno, che vengono suddivisi in sottointervalli elementari della durata di 5 minuti. Dopodiché, si individua l'intervallo della durata prefissata (suddiviso in sottoperiodi contigui di 5 minuti) cui corrisponde il massimo dell'altezza totale di pioggia. Quest'operazione viene eseguita fino a quando sono stati individuati nuovi intervalli della durata prestabilita all'interno dello stesso evento pluviometrico.

Il coefficiente di ragguglio areale è valutato come:

$$ARF(A, d) = \frac{h_A}{h_c(d)}$$

Con h_c massimo di precipitazione al pluviografo centrale e h_A massimo di precipitazione raggugliata.

Dallo studio di Moisello e Papiri è stata inoltre ricavata un'importante osservazione, relativa all'indipendenza del fattore di riduzione della precipitazione all'area dal tempo di ritorno T . Ne consegue l'eliminazione di questa variabile nell'elaborazione dei dati, e quindi nell'impiego dei risultati; in più, si può evitare di determinare la legge statistica secondo cui è distribuito il numero di intervalli di pioggia con ampiezza d , poiché il rapporto stimato per una determinata durata è lo stesso per tutti i tempi di ritorno. Infine, non ci sono condizioni che influenzino la scelta degli eventi di pioggia, poiché se il coefficiente è indipendente dal

tempo di ritorno T , allora non si devono prendere h_A e h_C con uguale frequenza probabile ma è sufficiente che siano tratte da serie di osservazioni corrispondenti, rispettando solo quello che si dice vincolo di simultaneità.

D'altra parte, come precedentemente riportato, diversi studi hanno riportato conclusioni contrastanti in merito alla dipendenza dell'ARF dal tempo di ritorno; in linea generale, non è sempre possibile trascurarne il legame e quindi si rende necessario ragionare in termini di rapporto tra altezze di pioggia con tempo di ritorno assegnata, individuate le distribuzioni di probabilità dei due massimi annuali per una precisa durata.

La formulazione ottenuta prevede:

$$ARF(A, d) = 1 - \exp(f_1 d^{f_2})$$

In cui:

$$f_1 = -2,472A^{-0,242}$$

$$f_2 = 0,6 - \exp(-0,6643A^{-0,235})$$

Dal confronto dei risultati stimati da Columbo, il cui studio è riferito alla stessa zona, sebbene ci si aspetti concordanza tra le valutazioni, queste sono profondamente diversi. I dati forniti da Columbo sono scarsamente dipendenti dalla durata, sovrastimati per durate inferiori a 2 ore e sottostimati per durate superiori a 4, 6, 12 ore. Questa divergenza può essere imputata alla diversa impostazione dei due studi, per cui Columbo propone di considerare i massimi su tutto il periodo, che giustifica la sovrastima e il calcolo dell'ARF con riferimento al centro di scroscio, piuttosto che al centro dell'area, il che giustificherebbe la sottostima. Infine, a causa dell'impossibilità di disporre di orologi perfettamente sincroni, Columbo ha adottato un metodo cautelativo che, come lui stesso ha riconosciuto, ha condotto indubbiamente a valutazioni in eccesso, almeno per le brevi durate.

Nel presente lavoro di tesi, la procedura applicata si basa sulla metodologia proposta dagli studi di Penta e sul Fiume Crati. Le restanti relazioni empiriche sono state impiegate al fine di poter calibrare e confrontare le performance dei diversi approcci empirici illustrati nell'interpolazione dei risultati.

Capitolo 2: Dataset meteo orario e giornaliero regionale, Eraclito e ERG5

2.1 Banca dati Eraclito ed ERG5

Una climatologia regionale ben nota tramite mappe climatiche locali accurate è di grande importanza per una gestione e una pianificazione sostenibili e appropriate del territorio e delle attività antropiche; e di interesse anche in settori quali l'agricoltura, l'industria, il trasporto. In Italia, la questione climatica è resa particolarmente sentita da fattori quali la complessità orografica e la crescente vulnerabilità di alcune regioni alla variabilità climatica, a causa dell'intenso aumento della densità di popolazione e al cambiamento significativo dell'uso del suolo, oltre che ai sempre più frequenti eventi e meteorologici estremi.

Gli strumenti climatici messi a disposizione da Arpae forniscono una rappresentazione della variabilità del clima attraverso mappe e grafici che evidenziano le deviazioni dalle medie climatiche e gli indici medi climatologici sulla Regione Emilia-Romagna. In questo contesto, il progetto Eraclito, finanziato dalla Regione Emilia-Romagna nell'ambito del Piano Telematico Regionale, nasce proprio per ottenere un'efficace descrizione del clima regionale, con un livello di dettaglio sufficiente ad avere un'adeguata riproduzione della variabilità spaziale e temporale.

L'Osservatorio Clima di Arpae rilascia, rispettivamente dal 2005 e dal 2012, dataset meteorologici, chiamati Eraclito ed ERG5, contenente dati climatici (Eraclito) e dati meteo-climatici (ERG5), delle principali variabili meteorologiche e agrometeorologiche e che coprono l'intero il territorio regionale ed un periodo che va complessivamente dal 1961 a oggi, con risoluzione temporale oraria e giornaliera.

Nel dettaglio, l'Osservatorio Clima produce la banca dati Eraclito, che si articola in una raccolta di dati giornalieri di precipitazioni e temperature (minima e massima) dal 1961 ad oggi, la cui versione corrente è la 4.4, rilasciata nel 2023. Il dataset costituisce la principale sorgente dati per la realizzazione dell'Atlante idroclimatico dell'Emilia-Romagna, strumento costituito da mappe annuali e stagionali relative a temperature, precipitazioni, evapotraspirazione potenziale e bilancio idroclimatico, indice di deficit o surplus di precipitazione, al netto dell'evapotraspirazione potenziale. L'Atlante fornisce una fotografia del clima in Emilia-Romagna e permette di analizzare, anche visivamente, l'entità del cambiamento climatico a livello regionale tramite il confronto dei valori medi delle variabili

rappresentate con riferimento al trentennio di riferimento climatico di base secondo le convenzioni dell'OMM (Organizzazione meteorologica mondiale, organismo delle Nazioni Unite), 1961-1990, e gli anni successivi, 1991-2015, porzione rilevante del passato trentennio climatologico, concluso nel 2020. Oltre alle tendenze termiche e pluviometriche in atto, sono riportate le proiezioni climatiche fino al 2050 in termini di aumento termico e variazione nei regimi delle piogge attese.

Il database ERG5 è un archivio realizzato dall'Osservatorio Clima e dall'Unità Territorio e reti di Arpa e Simc, con dati a passo giornaliero ed orario, di precipitazione dal 1991 al 2000 e di temperatura dell'aria, precipitazione, umidità relativa dell'aria, intensità (scalare e vettoriale) e direzione del vento, irradianza solare, bagnatura fogliare ed evapotraspirazione di riferimento, dal 2001 ad oggi. La versione attuale di quest'ultima è la 2.0, ed è stata rilasciata nel 2022. Queste lunghe serie storiche di dati mostrano che, negli ultimi sessant'anni, abbiamo sperimentato un aumento significativo sia delle temperature medie annuali, delle minime e delle massime, sia della frequenza degli eventi estremi in tutta la Regione.

I dataset sono ricavati a partire dai valori rilevati dalla rete delle stazioni meteorologiche, previo controllo di qualità in merito all'omogeneità statistica, alla sincronizzazione tra le misurazioni ed alla consistenza spaziale; tramite interpolazione spaziale su una griglia regolare costituita da 1131 celle di estensione di circa 25 km², che ricoprono l'intera Emilia Romagna (Figura 2.1). L'interpolazione consente di avere un archivio dati caratterizzato da:

- Completezza spaziale: in ogni cella con cui è discretizzato il territorio di competenza si ha sempre la disponibilità di dati, indipendentemente dalla presenza o meno di una stazione di misura.
- Completezza temporale: il malfunzionamento, le anomalie ecc. di uno o più strumenti non comporta l'assenza del dato, che non viene mai a mancare.
- Incrementalità: un aumento o una riduzione della densità dei dati o della rete non richiede un completo aggiornamento del sistema, che può essere adattato gradualmente, passo dopo passo.

Questi database costituiscono un sistema informativo georeferenziato webGis e continuamente aggiornato, a disposizione della comunità dei cittadini e degli enti locali.

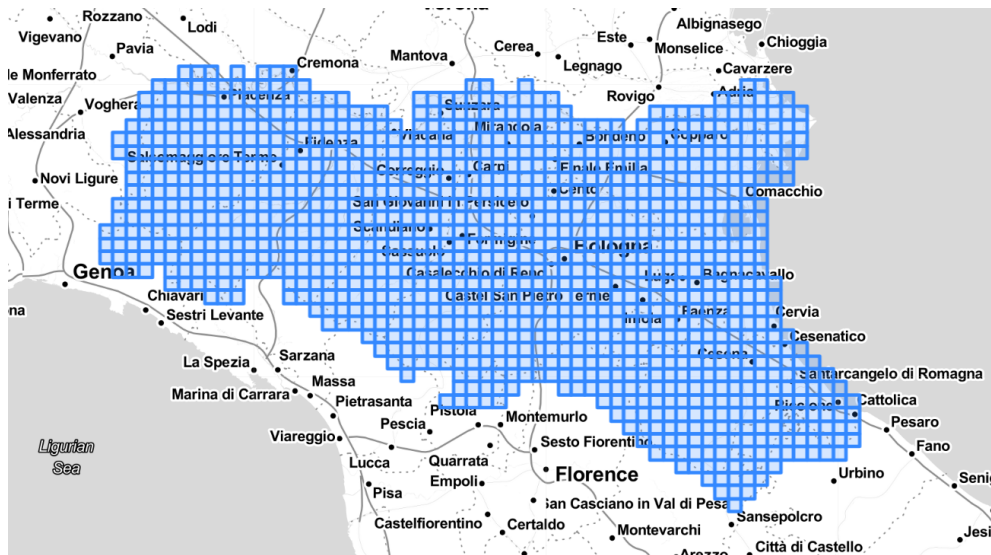


Figura 2.1 Interfaccia grafica per lo scaricamento dei dati ERG5_Eraclito (formato CSV) per cella e per anno.

2.2 La costruzione del dataset

I dataset attualmente disponibili sono il prodotto di un rigoroso controllo di qualità ed omogeneità delle serie storiche seguito dall'analisi climatica, giornaliera e oraria, su grigliato regolare ad alta risoluzione (circa 5 km) della precipitazione e altre variabili meteorologiche di interesse sul territorio dell'Emilia-Romagna, con elevata densità spaziale e temporale.

Si prosegue la trattazione relativa agli archivi Eraclito e ERG5 con riferimento all'elaborazione dei dati di precipitazione per la realizzazione del dataset Eraclito (v. 3.0), poiché analoga alle metodologie adottate per il dataset ERG5 ed ampiamente descritta nel numero 36 dell'International Journal of Climatology, *A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961-2010*, Antolini et al, 2016.

Le serie temporali di precipitazione giornaliera disponibili, provenienti da 254 località e raccolte nel periodo 1961-2010, sono state sottoposte ad una valutazione di qualità che si articola nelle seguenti operazioni:

- Eliminazione delle persistenze temporali spurie isolate;
- Controllo di omogeneità temporale tra le serie storiche;
- Controllo di sincronismo temporale a scala giornaliera tra le misurazioni;
- Verifica di plausibilità, coerenza interna tra le variabili e controlli climatologici;
- Controllo spaziale.

In particolare, le prime tre sono state eseguite una sola volta prima di procedere all'analisi, mentre le ultime due durante l'interpolazione.

Le misurazioni che superano i controlli preliminari sono fornite come input all'interpolazione spaziale su un modello digitale di elevazione ad alta risoluzione, e poi mediate, nel caso della versione del dataset esaminata, su una griglia irregolare triangolata con risoluzione variabile a seconda in base all'orografia locale. Questa scelta è legata all'elevato livello di dettaglio ottenuto in corrispondenza di montagne e colline, caratterizzate da una forte variabilità climatica spaziale, sebbene si abbia un'accuratezza ridotta sulle pianure, dove però il clima è più uniforme. La versione di Eraclito immediatamente successiva alla 3.0, la 3.1, è il risultato di un'interpolazione eseguita con le medesime tecniche e strategie, con la differenza che la griglia presa in considerazione è regolare, non più irregolare, con una risoluzione di circa 5 km, non più variabile. Nel dettaglio, viene impiegato il metodo di Shepard modificato, di seguito descritto.

L'errore legato all'interpolazione con riferimento alla precipitazione è mediamente una sottostima, pari fino al 25% per le precipitazioni intense e abbondanti nel semestre estivo.

2.3 La valutazione dei dati

L'affidabilità delle analisi climatiche è influenzata non solo dal buon funzionamento della rete strumentata, ma anche dalla qualità, quantità, omogeneità, sincronicità dei dati rilevati forniti in input. Per questo motivo è buona pratica sottoporre queste osservazioni, prima del loro utilizzo, ad opportune verifiche di qualità, che ne evidenzino gli errori e che consentano di eliminare i dati incoerenti e di individuare i periodi di mancato funzionamento non segnalato della stazione.

Sebbene siano attuati scrupolosi controlli di qualità, potrebbero esservi delle disomogeneità nelle serie di misurazioni causate da vari fattori, derivanti, ad esempio, dalle diverse fasi del processo di acquisizione e trattamento dei dati. I segnali climatici potrebbero essere distorti o irreparabilmente modificati da elementi come la posizione e la tipologia della stazione di misura, il suo spostamento o il cambio di osservatore o di uno specifico sensore; cause comuni di distorsioni che in genere si osservano con una improvvisa variazione nel valore medio della serie. D'altra parte, cambiamenti gradualmente legati ad esempio all'urbanizzazione o alla crescita della vegetazione in prossimità dello strumento, possono portare a modifiche della serie che si sviluppano nel tempo, creando dei trend alterati artificiali.

Infine, viene sviluppata un'ulteriore verifica sulla sincronicità della stazione dei dati osservati,

per l'individuazione di errori dovuti all'assegnazione della data di rilevamento delle singole misurazioni oppure di all'uso di diverse regole di calcolo del dato giornaliero, come nel caso di serie provenienti da stazioni automatiche, come riportato da Baker (1975), Schaal e Dale (1977), Janis (2002) e Holder et al. (2006). Questi errori possono comportare sfasamenti di uno o più giorni tra stazioni attigue, modificando la distribuzione spaziale della precipitazione.

2.3.1 I controlli di qualità

In sede di elaborazione dati, non è possibile trascurare la possibilità che questi siano affetti da errori, che ne possono alterare le caratteristiche statistiche. Lo scopo del controllo di qualità è proprio quello di individuare le misure affette da errore impiegando regole oggettive, così da eliminarle dall'analisi.

A partire dal 2001 (d.g.r. n. 2515) il Servizio Idro-Meteo-Clima di Arpa Emilia-Romagna (attuale Struttura Idro-Meteo-Clima di ARPAE) ha assunto il ruolo di gestore unitario delle reti di monitoraggio meteo-idro-pluviometrico (Figura 2.2) e, di conseguenza, di mantenere e alimentare le banche dati storiche. Il SIMC ha inoltre il compito di verificare la qualità dei dati meteo rilevati, per cui, in linea generale, i dati già presenti nel database ARPAE, presi in considerazione per la costruzione del dataset, superano quotidianamente i test di qualità di seguito riportati. Nel dettaglio, viene descritta la tecnica tratta da Pavan et al. (2003), con riferimento ai dati di precipitazione.

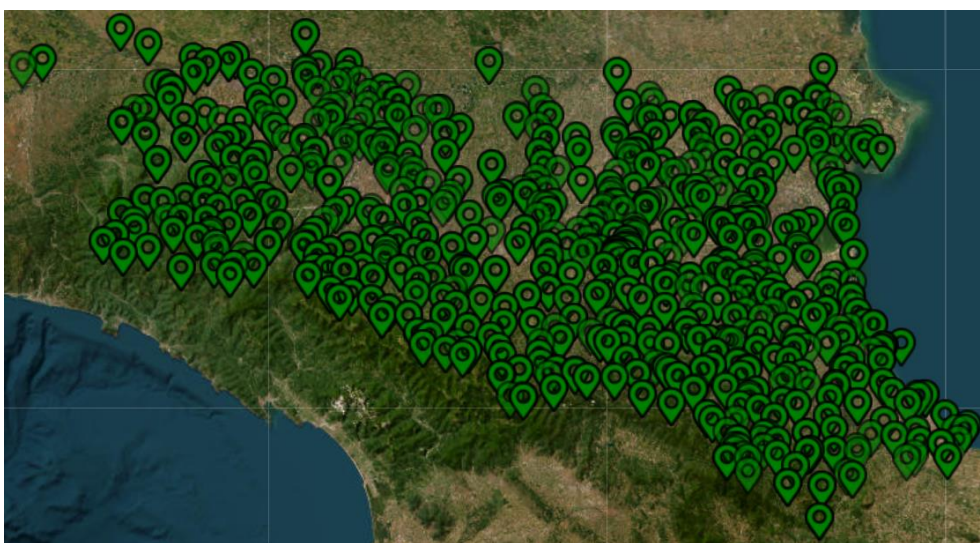


Figura 2.2 Rete di monitoraggio idro-meteo-pluviometrico regionale, risultato dell'integrazione di reti appartenenti a più enti operanti sul territorio con finalità differenti.

La condizione per l'applicazione del metodo è la disponibilità di una serie di osservazioni che ricopra un arco temporale di almeno 30 anni sia per la stazione di misurazione presa in esame che per alcune stazioni limitrofe; oppure che siano note le caratteristiche principali del clima della regione in cui si trova la stazione di interesse. Questo si pone l'obiettivo di identificare due tipologie di errori: i valori eccessivi di altezza di pioggia giornaliera non corrispondenti alla realtà, generati da errori nella digitazione o nella registrazione di valori di pioggia cumulata su più giorni; i periodi non segnalati di mancato funzionamento della stazione. In entrambi i casi il controllo si basa sul confronto della serie raccolta dalla stazione presa in esame con le serie rilevate da cinque stazioni in prossimità. Le stazioni circostanti vengono selezionate sulla base della quota, per via della complessa variazione della precipitazione con la stessa, in modo che: se la stazione si trova al di sotto dei 100 m s.l.m., vengono prese le stazioni che si trovano ad una quota che differisce con quella della stazione di interesse di meno di 200 m; mentre se si trova al di sopra di 1000 m s.l.m. vengono considerate le stazioni che cadono ad una quota che differisce da quella della stazione in esame di meno di 500 m. L'esito del controllo di qualità consiste nell'assegnazione a ciascun dato di un numero intero, flag, che può assumere valori quali:

- 0, dato valido;
- 1, dato sospetto, che può essere incluso nelle analisi;
- 2, dato molto probabilmente erroneo, da escludere dalle analisi.

I valori di soglia sulla base dei quali, nei seguenti controlli, si stabilisce l'affidabilità dei dati sono scelti in modo da minimizzare i dati della serie scartati dal controllo e massimizzare l'individuazione di possibili errori.

Dapprima viene applicato un gruppo di controlli logici, consistente nell'assegnazione di un flag pari a 2 (dato non valido) nel caso in cui il dato preso dalla stazione sottoposta a verifica sia negativo oppure nel caso in cui corrisponda al valore di dato mancante.

Si procede poi con l'eliminazione dei valori eccessivi non corrispondenti a quanto verificatosi in realtà, sulla base dei seguenti criteri:

- Se il valore della precipitazione è inferiore al dato climatologico o, rispetto al clima di quel giorno per quella stazione, presenta una differenza inferiore a 5 volte la deviazione standard, viene assegnato un flag nullo, ad indicazione della validità del dato.

- Se il valore della precipitazione è superiore al dato climatologico e, rispetto al clima di quel giorno per quella stazione, ha una differenza superiore a 5 volte la deviazione standard, il dato è considerato sospetto e viene indagato sulla base della stagione in cui ricade. In particolare, nel caso in cui il dato si collochi nella stagione estiva, dal 1° giugno al 31 agosto, il confronto viene svolto rispetto al valore climatologico medio calcolato prendendo in considerazione solo i giorni piovosi (altezza di pioggia giornaliera superiore o uguale a 1,0 mm) verificatisi nella decade in esame (per decade si intende sottoperiodo in cui viene suddiviso l'arco temporale mensile: dal 1° al 10, dall'11 al 20 e dal 20 all'ultimo del mese). Dunque, la modalità di assegnazione del flag è la seguente:

- Se la differenza tra il valore osservato e il valore medio ricavato è minore di 5 volte la deviazione standard, si assume che il dato in questione si riferisca ad un temporale intenso e viene ritenuto valido (flag pari a 0);
- Se tale differenza è maggiore di 5 volte la deviazione standard e in almeno in una delle stazioni limitrofe si osserva, nello stesso giorno, la registrazione di un'altezza di pioggia che differisce dal clima dei giorni piovosi della decade selezionata per più di una deviazione standard, si considera che il dato esaminato sia valido (flag pari a 0) poiché riferito ad un evento sinottico localmente molto intenso;
- Nei restanti casi si assume che il dato sia non valido (flag pari a 2).

Se il dato si colloca temporalmente nel rimanente periodo dell'anno, allora il confronto si esegue prendendo in esame i dati provenienti dalle stazioni vicine e corrispondenti cronologicamente a quello sottoposto a controllo:

- Se in almeno in una delle stazioni limitrofe si osserva un'altezza di pioggia che supera di 2 deviazioni standard il valore climatologico del giorno stesso, si assume che il dato sia valido (flag pari a 0);
- Diversamente il dato risulta invalido (flag pari a 2).

In ultima sede di verifica della serie raccolta dalla stazione in esame, si passa all'identificazione dei periodi di mancato funzionamento non segnalati. Condizione necessaria all'identificazione di questi periodi è l'assenza di pioggia per almeno 15 giorni consecutivi; è quindi indispensabile che non risulti, in tale periodo, alcun dato mancante e che i dati che si collocano nel suddetto non siano risultati non validi nelle operazioni di controllo antecedenti. In tal caso, si esamina l'affidabilità dei dati osservati in questo intervallo di tempo solo se

esistono almeno tre stazioni limitrofe (considerate tali in base alla quota, secondo il criterio precedentemente presentato) che abbiano registrato dati validi per i 15 giorni consecutivi in questione. Verificati questi requisiti, si calcolano la media dell'altezza di pioggia cumulata su questo periodo su tutte le stazioni in prossimità e il numero di giorni di pioggia (con un'altezza di pioggia cumulata sulle 24 ore maggiore di 1,0 mm) medio su tutte le stazioni nel periodo considerato. La valutazione del flag, ovvero dell'attendibilità del dato, si basa sui seguenti criteri:

- Se l'altezza di pioggia cumulata media sulle stazioni limitrofe è minore di 10,0 mm, i dati sono considerati validi, e il flag è impostato a 0;
- Se, viceversa, l'altezza di pioggia cumulata media sulle stazioni limitrofe è maggiore di 10,0 mm, si procede considerando il numero medio di giorni piovosi:
 - Se il numero medio di giorni piovosi è minore o uguale di 2, si considera che questo valore di precipitazione mediato sulle stazioni sia dovuto a un evento temporalesco, che può non aver coinvolto la stazione sotto verifica, il cui funzionamento nel periodo esaminato viene assunto corretto (flag pari a 0);
 - Se il numero medio di giorni piovosi è maggiore di 2, si assume che la stazione presa in considerazione non abbia funzionato correttamente e quindi si prendono come non validi (flag pari a 2) i dati registrati dalla stazione in questi 15 giorni.
- Infine, nel caso in cui non sia possibile individuare almeno tre stazioni limitrofe disponibili con le quali eseguire il confronto, le osservazioni riportate dalla strumentazione indagata sono considerati sospetti (flag unitario).

Nel momento in cui l'esito della verifica sia l'identificazione di un periodo di mancato di funzionamento di 15 giorni, e quindi di 15 valori di altezza di pioggia giornaliera non validi, ogni valore nullo di precipitazione registrato successivamente viene preso come non valido anch'esso, finché non si verifica il primo giorno di registrazione di un'altezza di pioggia non nulla, compreso nel periodo di non validità, dal momento che probabilmente si tratta di un valore di pioggia cumulata non segnalato.

Al termine di questa procedura, si assume che gli errori ricercati, in particolare valori eccessivi e persistenze temporali isolate spurie, siano stati eliminati.

Nelle operazioni di controllo per filtrare i dati ritenuti idonei dal punto di vista della qualità per la costruzione del dataset, sono implementate delle verifiche di particolare importanza per

i dati estratti dalle singole stazioni, non sottoposti ai controlli di qualità del SIMC. In dettaglio, vengono identificate e eliminate le persistenze temporali isolate spurie di precipitazioni nulle (superiori a 15 giorni) mediante il confronto con diverse località circostanti. Sono stati utilizzati, inoltre, altri test, quali controlli di plausibilità e coerenza interna tra le variabili e controlli climatologici temporali e spaziali, di seguito descritti, anche più complessi dei test di qualità precedentemente trattati.

2.3.2I controlli di omogeneità

Una serie temporale omogenea si definisce come tale nel caso in cui la variabilità del segnale dipende esclusivamente da fattori climatici. Le serie di dati ottenute, al netto dei controlli di qualità, possono essere in ogni caso falsate o condizionate da fattori esterni, anche antropici, che intervengono in sede di acquisizione ed elaborazione dati, alterando quindi la valutazione del trend effettivo. Le osservazioni raccolte possono essere influenzate da un sistema complesso di elementi, quali la tipologia e la posizione della stazione, al territorio in cui si colloca, alle tecniche di rilevamento impiegate dall'operatore e dalle formule scelte per la derivazione di alcune grandezze. Modifiche o sostituzione della strumentazione; lo spostamento della stessa; un cambio di osservatore o l'impiego di nuove tecniche di elaborazione delle misure sono tra le cause più comuni dell'introduzione di disomogeneità nella serie raccolta, corrispondenti a brusche variazioni del valore medio della serie. D'altra parte, modifiche gradualmente dell'ambiente circostante alla stazione, che comportano un cambiamento della sua esposizione, legati all'urbanizzazione o alla crescita della vegetazione nell'area in prossimità, generano sulla serie un effetto che si protrae nel tempo, dando luogo a trend artificiali.

Non esiste una serie di misurazioni che può essere ritenuta omogenea a priori e dal momento che l'impiego di una serie temporale omogenea con variazioni del segnale climatico riconducibili a cause di natura meteo-climatica è di grande importanza per le successive analisi climatiche, a tal merito, sono state sviluppate diverse metodologie e tecniche per l'individuazione di segnali di natura non climatica all'interno delle serie esaminate. Genericamente, si possono adottare due possibili metodi per l'identificazione delle disomogeneità all'interno di una serie, ovvero i metodi diretti e i metodi indiretti. I primi si basano sull'analisi dei metadati, informazioni che documentano la storia della stazione, che permettono di determinare con precisione se e quando è avvenuta la discontinuità e che cosa

l'ha causata. L'adozione di un metodo diretto, se da un lato comporta il riconoscimento della discontinuità e la sua eliminazione sulla base di informazioni dettagliate ed affidabili; dall'altro è influenzato dalla qualità dei metadati, dalla loro completezza e dalle tecniche usate per quantificare gli eventuali fattori di correzione: raramente è possibile l'estrazione dai metadati dei valori correttivi da utilizzare per eliminare la discontinuità.

Più diffusi, per via della loro facilità di applicazione e della comune indisponibilità dei metadati, sono i metodi indiretti. Questi consistono nell'esecuzione di test statistici che, con il supporto dei metadati, laddove possibile, individuano una possibile discontinuità nella serie di precipitazione e forniscono un metodo semplice per ottenere i valori correttivi da impiegare, in funzione della costruzione di una serie temporale da assumere come riferimento per la valutazione della serie candidata relativa alla stazione in esame. La scelta della serie di riferimento è una procedura delicata, poiché si basa sulla scelta di un determinato numero di stazioni in prossimità statisticamente omogenee che permettano di realizzare una serie rappresentativa della variabilità climatica dell'area in cui si colloca la stazione esaminata. Generalmente, le stazioni considerate sono scelte guardando ai seguenti parametri: distanza e differenza di quota rispetto alla stazione candidata; coefficiente di correlazione fra le serie di valori medi annuali o stagionali, o cumulate annuali o stagionali.

Due diversi metodi indiretti applicati a una stessa serie di dati possono produrre risultati differenti, se non addirittura discordanti, per cui solitamente si prendono in esame i risultati di più tecniche di analisi, con il supporto e la conferma dei metadati se disponibili.

L'omogeneizzazione si complica ulteriormente quando si passa dalla scala annuale e mensile a quella giornaliera: l'elevata variabilità che caratterizza le serie dei dati giornalieri può produrre nella serie discontinuità climatiche difficilmente distinguibili da discontinuità artificiali. L'obiettivo dell'identificazione delle breakpoint, d'altra parte, si accompagna al rischio di operare un overadjustment, ovvero di eccessiva o errata correzione della serie. A tal riguardo, si può svolgere una verifica a scala annuale assumendo che serie prive di breakpoint a livello annuale possano essere considerate omogenee anche a livello giornaliero oppure, ragionando su scala giornaliera, si possono escludere le serie che presentano evidenti discontinuità artificiali o, in alternativa, considerarne solo la parte omogenea, come suggerisce l'OMM.

L'attuazione di un progressivo recupero e digitalizzazione dei dati disponibili sul territorio regionale, operato dal Servizio Idro-Meteo-Clima, ha consentito di avviare, sin da prima delle prime realizzazioni del dataset, il test di omogeneità SNHT (Standard Normal Homogeneity

Test) (Alexandersson e Moberg, 1997) sulle serie storiche di temperatura e di precipitazione. Nell'ambito della procedura di realizzazione del dataset Eraclito descritta nell'International Journal of Climatology, ciascuna serie temporale di osservazioni è stata verificata per l'omogeneità temporale rispetto ad altre serie di stazioni vicine, con riferimento ai soli dati annuali (Alexandersson e Moberg, 1997; Vincent, 1998). Sono stati applicati tre diversi metodi statistici: il test di omogeneità SNHT è il test principale utilizzato per verificare l'omogeneità della serie temporale (Alexandersson e Moberg, 1997); i test di Craddock (1979) e Vincent (1998) vengono utilizzati a supporto del SNHT. Solo le serie temporali statisticamente omogenee sono state utilizzate come input per l'analisi; nel dettaglio, con riferimento alla precipitazione, hanno superato i test di omogeneità 254 serie temporali.

2.3.2.1 Standard Normal Homogeneity Test

Il test SNHT (Alexandersson and Moberg, 1997; Tuomenvirta, 2001) è uno dei più comuni in bibliografia e si basa sul confronto fra la serie candidata ed una serie di riferimento, costruita come la media pesata delle serie provenienti da stazioni vicine selezionate, caratterizzate da alti valori di correlazione con la serie candidata, utilizzando come pesi i coefficienti di correlazione delle stazioni scelte con la serie da controllare. Nel caso di serie di precipitazione, il test è basato sull'assunzione che il rapporto o la differenza tra la stazione da controllare e la serie di riferimento è quasi costante nel tempo.

Indicando con $\{Y_i\}$ la serie temporale candidata e con $\{X_i\}$ la serie di riferimento costruita, si ricava la serie dei rapporti $\{Q_i\}$:

$$Q_i = \frac{Y_i}{\frac{\sum_{j=1}^k \rho_j^2 X_{ji} \bar{Y}}{\bar{X}_j} \frac{1}{\sum_{j=1}^k \rho_j^2}}$$

Dove:

- k è il numero delle stazioni scelte per creare la serie della stazione di riferimento;
- j è l'indice della stazione di riferimento;
- i è l'indice del time-step (anno, stagione, mese, giorno; in questo caso l'anno considerato);
- X_{ji} è il valore all'istante i -esimo della j -esima stazione;
- \bar{X}_j è il valore medio della j -esima serie;
- Y_i è il valore all'istante i -esimo della stazione da testare;

- \bar{Y} è il valore medio della serie candidata;
- ρ_j^2 è il quadrato del coefficiente di correlazione tra la serie candidata Y e la serie di riferimento X, che corrisponde al peso relativo ad ogni serie di riferimento scelta.

Standardizzando la serie dei rapporti $\{Q_i\}$:

$$Z_i = \frac{Q_i - \bar{Q}}{\sigma_Q}$$

Dove σ_Q è la deviazione standard e \bar{Q} il valore medio della serie; si crea la serie $\{Z_i\}$, che si assume sia descritta da una distribuzione normale $N(\mu, \sigma)$. Con l'obiettivo di individuare un singolo shift del valore medio della serie candidata, le due ipotesi sono le seguenti:

- Ipotesi nulla, che afferma che la serie in esame è omogenea:

$$H_0: Z_i \in N(0,1); i \in \{1, \dots, n\}$$

Se la serie è omogenea, la serie $\{Z_i\}$ avrà media nulla e deviazione standard unitaria.

- Ipotesi alternativa, che indica che il livello medio della serie cambia bruscamente nell'anno a:

$$H_1: \begin{cases} Z_i \in N(\mu_1, 1); i \in \{1, \dots, a\} \\ Z_i \in N(\mu_2, 1); i \in \{a + 1, \dots, n\} \end{cases}$$

Se si verifica una disomogeneità nell'anno a, $\{Z_i\}$ avrà una media μ_1 per i che va da 1 ad a e una media μ_2 per i che va da a+1 ad n.

La statistica T è calcolata con la seguente relazione:

$$T_{max}^S = \max_{1 \leq a \leq n-1} \{T_a^S\} = \max_{1 \leq a \leq n-1} \{a\bar{z}_1^2 + (n-a)\bar{z}_2^2\}$$

Dove \bar{z}_1 e \bar{z}_2 sono i valori medi della serie $\{Z_i\}$ prima e dopo lo shift dell'anno a. L'ipotesi nulla non è convalidata nel caso in cui il massimo della T_a^S è sopra il valore di soglia, funzione della lunghezza della serie n e dal livello di confidenza scelto (90% o 95%, Tabella 2.1).

n	10	20	30	40	50	70	100
T90	5,05	6,10	6,65	7,00	7,25	7,55	7,85
T50	5,70	6,95	7,65	8,10	8,45	8,80	9,15

Tabella 2.1 Valori soglia di T in funzione della lunghezza della serie in esame n e del livello di confidenza scelto.

Questo test ha il vantaggio di fornire un risultato oggettivo su un salto del segnale climatico che si traduce in uno shift del livello medio a partire da uno specifico istante (anno). Uno dei pochi svantaggi del SNHT è la scarsa performance nell'individuazione di discontinuità che si verificano ai margini, inizio e fine, della serie.

2.3.2.2 Test di Craddock

Il test di Craddock (Craddock, 1979) è un test non parametrico matematicamente più semplice del test SNHT e si basa sull'andamento nel tempo della quantità S_n , ricavata utilizzando una o più serie di riferimento $\{Y_i\}$ ed una serie candidata $\{X_i\}$. Sostanzialmente, questo test, accumula le differenze fra la serie candidata e la serie di riferimento. Nel caso di serie di precipitazione, si definisce:

$$c = \frac{\bar{y}}{\bar{x}}$$

E:

$$d_i = cx_i - y_i$$

Dove:

- i è l'indice del time-step (annuale);
- \bar{x} è il valore medio della serie da testare;
- \bar{y} è il valore medio della serie di riferimento.

Si ricava la serie S_n come:

$$S_n = \sum_{i=1}^n d_i$$

Nel caso in cui i valori della serie candidata $\{X_i\}$ siano sovrastimati o sottostimati rispetto alla serie di riferimento $\{Y_i\}$, la serie S_n è caratterizzata rispettivamente da un trend positivo o negativo. Lo scopo è quello di individuare le disomogeneità, e viene raggiunto indagando la presenza di cambiamenti della pendenza media della S_n , ovvero di discontinuità nella derivata prima. Nel caso in cui la serie candidata $\{X_i\}$ sia omogenea, la serie S_n fluttua attorno al valore nullo.

Il test di Craddock risulta più sensibile ad eventuali valori estremi isolati nella serie. Tramite il confronto tra la serie candidata con ognuna delle serie di riferimento, consente di indagare anche l'efficacia della scelta delle stazioni di riferimento stesse, così da rivedere la selezione se una o più delle serie di riferimento prese non è adatta al confronto con la serie da verificare.

2.3.2.3 Test di Vincent

Il test di Vincent (Vincent, 1998; Vincent and Gullett, 1999) consiste nel calcolo di alcuni modelli di regressione lineare fra la serie candidata e la serie di riferimento, costruita come la media aritmetica delle serie delle stazioni di riferimento. In particolare, se si utilizza una sola serie di riferimento la regressione sarà semplice, nel caso in cui si abbiano più serie di riferimento allora la regressione sarà multipla.

Nel caso di regressione semplice, si applicano successivamente dei modelli regressivi secondo lo schema:

- Modello 1: il primo modello fornisce come esito l'indicazione dell'eventuale omogeneità o disomogeneità della serie candidata:

$$y_i = a + cx_i + e_i$$

Dove:

- y_i è la serie candidata;
- a e c sono i parametri della regressione (stimati attraverso il metodo dei minimi quadrati);
- e_i sono i residui della regressione, ovvero le differenze fra i valori della serie candidata e quelli fittati dati dal modello.

Il modello può essere testato, ad esempio, con una statistica di Durbin-Watson, D :

$$D = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Questo viene confrontato con dei valori di riferimento superiori e inferiori, rispettivamente d_u e d_l , per cui se D è superiore a d_u la serie è omogenea e il test viene fermato; se è inferiore a d_l allora la serie candidata potrebbe non essere omogenea e il test prosegue; se è compreso tra d_l e d_u non è possibile trarre conclusioni e interrompere il test.

Un altro metodo che può essere impiegato per stabilire l'effettiva presenza di breakpoint è illustrato da Chatfield (1984), che consiste nel generare una serie attraverso l'autocorrelazione dei residui a diversi lag k :

$$r_k = \frac{\sum_{i=1}^{n-k} (e_i - \bar{e})(e_{i+k} - \bar{e})}{\sum_{i=1}^n (e_i - \bar{e})^2}$$

Nel caso in cui i coefficienti di autocorrelazione, specialmente a lag bassi, non siano significativamente diversi dal valore nullo, la serie candidata è omogenea; altrimenti, autocorrelazioni significative per lag consecutivi evidenziano la presenza di discontinuità o di trend.

- Modello 2: nel caso in cui dal modello 1 risulti un possibile trend, se ne verifica la significatività tramite il modello 2:

$$y_i = a_2 + b_2i + c_2\bar{x} + e_i$$

Dove viene introdotta la variabile b_2 , pendenza della retta di regressione nel tempo.

Tramite il test di Mann-Kendall ¹ (con soglia di significatività 0.95), viene stabilita la significatività del trend.

- Modello 3: nel caso in cui dal modello 1 risulti una possibile discontinuità, se ne verifica la significatività tramite il modello 3:

$$y_i = a_3 + b_3I + c_3\bar{x} + e_i$$

Dove il parametro b_3 rappresenta l'entità dello step. Vengono presi in considerazione soltanto gli step dalla posizione i pari a 4 ad $n-3$.

La significatività dello step viene stabilita tramite il test F, per cui si confrontano i modelli 1 e 3 e, quando sono fittati al dataset, si calcolano la somma dei quadrati dei residui $RSS1$ ed $RSS3$, e quindi la statistica F:

$$F = \frac{\frac{RSS1 - RSS3}{DF1 - DF3}}{\frac{RSS3}{DF3}}$$

Dove:

- $DF1$ pari a $n-2$;
- $DF3$ pari a $n-3$.

Il test è più complesso dei precedenti test di omogeneità e consente di stabilire se un'eventuale disomogeneità della serie è causata da un trend lineare o da una improvvisa variazione della media. Per questo motivo è un buon supporto all'interpretazione dei risultati degli altri test, soprattutto nel caso di serie dubbie.

2.3.3I controlli di sincronicità

La sincronia dei dati è essenziale per la produzione di mappe di analisi giornaliere; di conseguenza, si rende necessaria l'esecuzione di un'analisi della sincronicità dei dati giornalieri di una stazione di precipitazione, in modo da individuare eventuali incongruenze

¹ Il test non-parametrico di Mann-Kendall (Mann, 1945; Kendall, 1975) ha l'obiettivo di individuare la significatività di un trend in una serie temporale di dati ed è basato sul concetto di "ranking", cioè sulle caratteristiche di ordinamento della serie storica in esame.

dovute, per esempio, a differenti finestre di rilevazione a scala giornaliera presso le diverse stazioni.

I dati utilizzati per la costruzione del dataset hanno una finestra temporale di osservazione che va dalle 9 del mattino, 09:00 LST (UTC+1), alle 9 del mattino seguente, 09:00 LST (UTC+1), Local Sidereal Time (LST), e sono convenzionalmente attribuiti al giorno in cui termina il periodo di osservazione. È però possibile che i dati registrati siano attribuiti al giorno in cui è iniziato il periodo di osservazione; attribuzione erronea temporale che può causare problemi nelle mappe di analisi giornaliera perché vengono interpolati dati che si riferiscono a giorni diversi (Haylock et al., 2008). Nonostante ci sia coordinazione tra diversi uffici nella convalida dei dati, non è insolito trovare differenze tra le procedure di segnalazione nei bollettini ufficiali. Inoltre, sono state spesso rilevate asincronie tra le stazioni nei set di dati che contengono dati giornalieri registrati sia automaticamente (per i quali il periodo di cumolazione in molti casi è 00:00-24:00) che manualmente (09:00-09:00), come riportato da Baker (1975), Schaal and Dale (1977), Karl et al. (1986), Janis (2002), Belcher and De Gaetano (2003) e Holder et al. (2006). Questo non è il caso dei dati giornalieri di precipitazione utilizzati in questo lavoro.

Per valutare la presenza di shift temporali spurii tra le stazioni e correggerli, è stata definita e applicata una procedura basata sulla costruzione di serie temporali di valori di correlazione annuali tra la stazione di test e una stazione di riferimento. Le correlazioni annuali vengono calcolate assumendo sia che tra le due non ci sia alcun ritardo e quindi che ci sia completa sincronicità; che assumendo che la serie temporale di test abbia un ritardo di 1 o 2 giorni in una delle due direzioni temporali, cioè prima o dopo la stazione di riferimento. Le serie temporali di valori di correlazione con e senza ritardo vengono confrontate. Quando le serie temporali sono sincrone, la correlazione per il ritardo 0 dovrebbe sempre presentare valori più elevati rispetto a tutte le altre correlazioni ritardate. In caso contrario, è presente un'asincronia nella stazione di test. In altre parole, il confronto tra le varie serie per mette di identificare la fase più probabile tra le due serie, dal momento che il valore della correlazione sarà massimo per la serie di correlazioni calcolate con lo sfasamento corretto, inferiore negli altri casi considerati.

Individuati i periodi di sfasamento tra due stazioni, se ne identificano le date esatte di inizio e di fine sfasamento confrontando i grafici della media mobile, su diversi filtri temporali (10 giorni, 50 giorni), della serie giornaliera delle differenze assolute tra i dati giornalieri della serie temporale di test e i valori interpolati nella stessa posizione ottenuti utilizzando solo le

stazioni circonvicine. L'utilizzo della media mobile permette di eliminare il rumore delle serie di dati giornalieri, evidenziando il segnale di sfasamento. Come prima, i grafici vengono calcolati più volte, assumendo diversi valori di sfasamento tra le due serie, per poi confrontare gli andamenti ottenuti.

Una volta identificate le date esatte di inizio e fine del periodo di asincronia, tutti i dati della serie temporale di test appartenenti a tale periodo vengono spostati di conseguenza.

2.3.4 Altri controlli

Tutte le variabili climatiche sono state sottoposte a controlli di plausibilità e coerenza interna tra le variabili, così come sono stati applicati controlli climatologici, temporali e spaziali.

- Il controllo di plausibilità assicura che ciascun valore osservato rientri nell'intervallo di osservazione strumentale.
- Il controllo interno garantisce la coerenza tra diverse variabili meteorologiche osservate in una particolare località (ad esempio, la temperatura minima giornaliera deve essere sempre inferiore o uguale alla temperatura massima).
- I controlli climatologici verificano se i dati giornalieri in una stazione particolare rientrano in un intervallo accettabile di valori climatologici, fissato sulla base delle osservazioni locali passate.
- Il controllo temporale confronta il tasso di variazione di una variabile con la sua distribuzione statistica, escludendo variazioni improvvise particolarmente intense.
- Infine, il controllo spaziale verifica se il valore osservato in una stazione è coerente dal punto di vista spaziale con quelli osservati nello stesso momento in un gruppo di stazioni vicine. A tale scopo, è stato utilizzato, per ogni variabile, un approccio di convalida incrociata a due fasi:
 - La prima fase ha l'obiettivo di individuare le stazioni in cui i residui di convalida incrociata appaiono superiori a una specifica soglia. I residui vengono calcolati sfruttando le informazioni tratte dai dati delle stazioni circonvicine, inclusa la loro deviazione standard, la variabilità dell'altitudine e la distanza relativa media.
 - Nella seconda fase, ciascun dato viene nuovamente verificato, utilizzando solo i dati delle stazioni vicine che hanno superato la prima fase come riferimento, evitando che dati sospetti possano influenzare il controllo.

Al netto dei controlli sopra descritti, sono stati impiegati nell'analisi climatologica, i dati di precipitazione, e di temperatura, provenienti dalle stazioni evidenziate nella Figura 2.3.

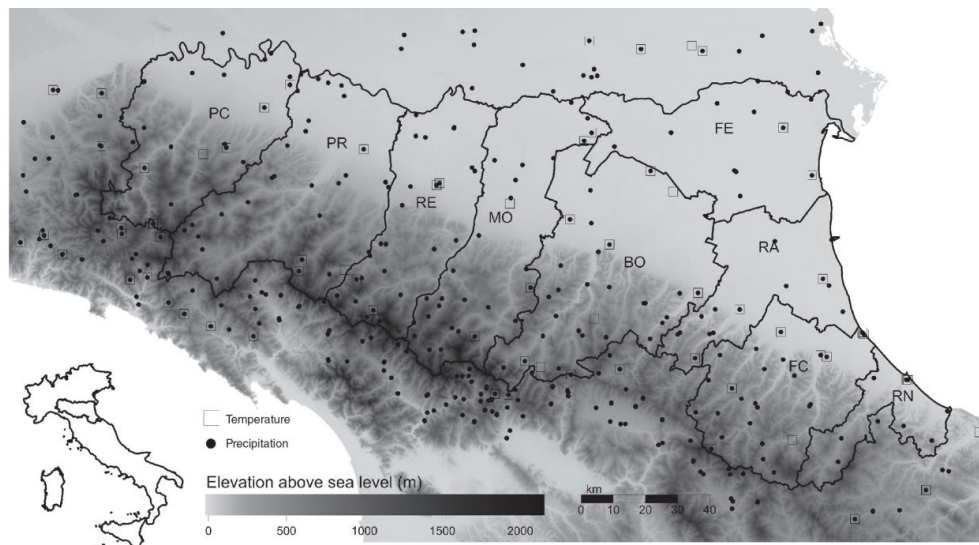


Figura 2.3 Mappa della regione Emilia-Romagna (Italia), province amministrative (con codici) e posizione delle stazioni meteorologiche osservative di precipitazione e temperatura incluse nell'analisi al netto dei controlli (Antolini et al.).

La versione successiva del dataset, la 3.1, rilasciata nel 2017, è realizzata sulla base delle seguenti stazioni termometriche e pluviometriche (Figura 2.4).

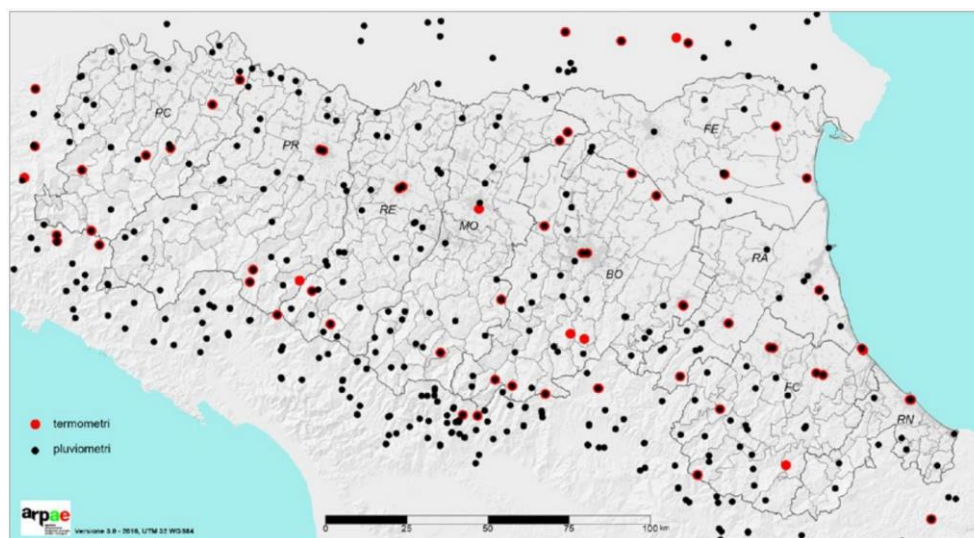


Figura 2.4 Stazioni termometriche e idrometriche impiegate per la costruzione del dataset Eraclito v. 3.1.

Si può osservare un incremento delle stazioni di misura considerate nell'analisi, con positive conseguenze sui risultati dell'interpolazione.

2.4 L'interpolazione

L'interpolazione spaziale è un metodo che consente di prevedere i valori assunti da una variabile in posizioni non campionate, utilizzando i dati provenienti dalle località campionate circostanti.

2.4.1 La griglia

Prima procedere con l'interpolazione dei dati, è necessario stabilire il passo di ricostruzione del dato in superficie, cioè la risoluzione e quindi il tipo di griglia.

La scelta della risoluzione migliore da adoperare dipende solitamente dalla densità dei dati noti (distanza media fra i vari punti) e dall'uso che si intende fare dei dati interpolati.

In generale, in climatologia, le stime vengono calcolate come medie spaziali sulle celle di una griglia regolare, come nel caso dei dataset correnti Eraclito e ERG5. D'altra parte, la variabilità spaziale, specialmente per la temperatura, è più elevata nelle aree dove l'orografia è più complessa, mentre per le precipitazioni può essere rilevante nelle stagioni più calde, quando le precipitazioni possono risultare da processi convettivi.

Le griglie regolari permettono una ricostruzione regolare a celle quadrate. Il valore associato ad ogni cella è ottenuto come la media fra tutti i valori interpolati per ogni punto disponibile sul Digital Elevation Model (DEM) che cade all'interno di questa cella. Il DEM attualmente utilizzato per l'Emilia-Romagna ha una risoluzione di 500 m.

L'analisi presentata è calcolata su una rete triangolata irregolare (Triangulated Irregularly spaced Network, TIN), in cui la risoluzione orizzontale è proporzionale alla variabilità orografica. Le TIN sono strutture spaziali basate su vettori, composte da celle triangolari contigue e non sovrapposte di dimensioni variabili. Questo approccio consente di ottenere il massimo dettaglio solo nelle aree in cui è necessario, con maglie più fitte nelle zone collinari e montane rispetto alle zone di pianura; riducendo al minimo lo spazio di archiviazione e il tempo computazionale. La TIN viene creata tramite l'algoritmo di triangolazione Delaunay (Li et al., 2005), utilizzando un set di dati con risoluzione orizzontale di 90 metri proveniente dalla NASA Shuttle Radar Topographic Mission. In questo caso i criteri di definizione dei triangoli sono stati modificati rispetto all'algoritmo standard, riducendo al minimo la variabilità orografica all'interno dei triangoli, anziché le differenze di elevazione tra il piano del triangolo e le celle del Modello di Elevazione Digitale (DEM). In questo modo, i triangoli risultano avere un'altitudine più uniforme rispetto alle TIN standard e alle griglie quadrate ordinarie.

La TIN regionale risultante è composta da 4302 triangoli e 2301 vertici, con una risoluzione equivalente massima (dimensione della cella quadrata con la stessa area) di circa 1300 m nelle zone montuose e una risoluzione equivalente minima di circa 4200 m nella pianura. La copertura della TIN si estende oltre i confini amministrativi regionali al fine di includere completamente tutti i bacini fluviali dell'Emilia-Romagna monitorati dall'ARPA-SIMC (Figura 2.5).

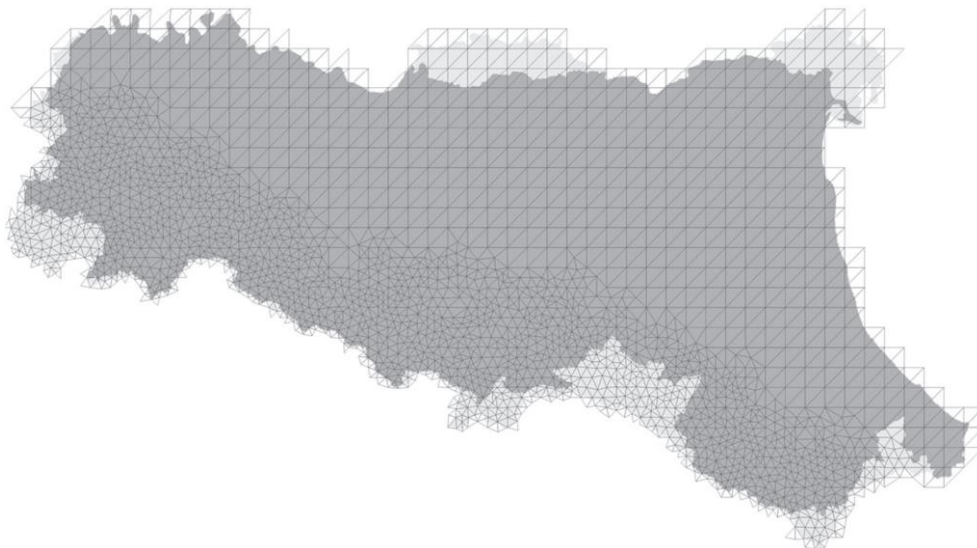


Figura 2.5 Griglia triangolata irregolare utilizzata nell'analisi, sull'area amministrativa dell'Emilia-Romagna (grigio scuro) e sull'area del bacino fluviale monitorata da ARPA-SIMC (grigio chiaro).

Oltre a ridurre al minimo lo spazio di archiviazione e il tempo di elaborazione, le TIN offrono un altro vantaggio: la possibilità di campionare i dati su griglie regolari (assegnando un peso a ciascun triangolo in base all'area intersecata da ciascuna cella regolare), con risoluzioni adatte alle applicazioni dell'utente, che può essere scelta nell'intervallo tra la risoluzione minima e massima della TIN.

La procedura di analisi inizia con una prima interpolazione dei dati giornalieri su una griglia regolare del Modello di Elevazione Digitale (DEM) a 500 metri (aggregata dall'originale DEM a 90 metri). I dati interpolati vengono quindi mediati sulle celle della TIN, al fine di ottenere un completo insieme di dati grigliati. In questo caso, la scelta è stata quella di impiegare un'unica struttura TIN per tutte le variabili, consentendo in questo modo di calcolare indici di combinazione che siano funzione di tutte o alcune delle variabili disponibili.

2.4.2 Metodo di gridding e distanza topografica

Il metodo di interpolazione che è stato adottato su maglia triangolata irregolare e di seguito descritto è lo stesso impiegato per la realizzazione delle successive versioni dei dataset.

I dati di precipitazione vengono interpolati utilizzando un algoritmo modificato di Shepard (1968), basato su uno schema di ponderazione in cui i pesi vengono calcolati inversamente alla distanza e alla direzione rispetto al punto considerato (inverse distance and direction weighting). Uno dei suoi vantaggi è la dipendenza direzionale dei pesi, che consente alle stazioni isolate di avere un peso maggiore rispetto a gruppi di stazioni. In altre parole, i punti raggruppati in aree a maggior densità di valori osservati contribuiscono con un peso minore alla somma pesata.

L'Inverse Distance Weighted (IDW) è una tecnica largamente utilizzata per la modellazione di superfici. Si tratta di un metodo deterministico che mette in atto un'interpolazione multi variabile con un set di valori noti e che si basa sull'assunzione che le misure rilevate in punti vicini tendono ad essere più simili rispetto al caso in cui i punti siano lontani tra loro. In linea generale, la funzione di interpolazione IDW per la stima di un valore interpolato Z in un punto di coordinate (x, y) sulla base delle misure $Z(x_i, y_i)$, per $i=0, 1, \dots, N$, consente di assegnare il valore del punto sconosciuto Z attraverso la media pesata dei punti noti Z_i , tramite la seguente relazione generale:

$$Z(x, y) = \sum_{i=1}^N \lambda_i Z(x_i, y_i)$$

Dove:

- $Z(x, y)$ è il valore stimato nel punto arbitrario non campionato interpolato, di coordinate (x, y) ;
- $Z(x_i, y_i)$ è il valore nel punto conosciuto, di coordinate (x_i, y_i) ;
- d è la distanza tra il punto interpolato x ed il punto conosciuto x_j ;
- N è il numero totale di punti campionati usati per l'interpolazione;
- λ_i è il valore della funzione peso λ , associato al punto i -esimo e inversamente proporzionale alla distanza, d_i , dal punto sconosciuto al punto noto.

Il peso, generalmente, è determinato con una relazione del tipo:

$$\lambda_i = w(d_i)$$

Ed è inversamente proporzionale alla distanza tra il punto non campionato all'i-esimo punto campionato, per cui:

$$\lim_{d \rightarrow \infty} w(d) = 0$$

L'andamento della funzione $w(d)$ può essere di due tipi:

- Potenze negative:

$$w(d_i) = d^r$$

- Esponenziale negativo:

$$w(d_i) = e^{rd}$$

Oppure

$$w(d_i) = e^{rd^2}$$

Dove r è l'esponente proprio della funzione peso.

La IDW assume che ogni misura abbia un'influenza limitata che decresce con la distanza dal punto di misurazione, assegnando ai dati utilizzati nell'interpolazione un peso che diminuisce man mano che ci si allontana dal punto incognito. Si osserva, infatti, che i pesi diminuiscono con la distanza e che un incremento dell'esponente r applicato a d o all'esponenziale comporta un incremento dell'influenza del valore della distanza sul risultato: se l'esponente è 0, tutti i punti entro il raggio di ricerca sono pesati ugualmente e si ricade nel caso della media mobile semplice; aumentandolo, si assegna un peso maggiore ai punti interpolati ad una ridotta distanza dai punti noti. Se l'esponente tende ad infinito, il peso viene attribuito interamente all'osservazione più prossima al punto interpolato.

Questo tipo di approccio è adatto alla rappresentazione in zone dove si hanno brusche variazioni superficiali (zone a bassa correlazione spaziale), poiché non tiene conto delle variazioni superficiali graduali e in generale del trend della superficie. La variabilità locale potrebbe però essere influenzata da fenomeni come il sollevamento orografico, che si realizza quando una massa d'aria atmosferica viene forzata a spostarsi da un punto a minor elevazione ad uno a quota maggiore, a causa della particolare conformazione orografica del terreno; o l'ombra pluviometrica, anche definita effetto rainshadow, un'area resa desertica dall'assenza di precipitazioni, favorevoli alla crescita di vegetazione, a causa di catene montuose che

bloccano tutti i sistemi meteorologici piovosi. Questi fattori possono amplificare la variabilità spaziale locale o ostacolare l'influenza di anomalie locali su aree più vaste. Per questo motivo, lo schema di Shepard è stato modificato, generalizzando la distanza cartesiana, d_i , tra i punti in una distanza topografica, D_t : la formulazione applicata per la costruzione del dataset è derivata da Antolini e Tomei (2006), con modifiche tratte dalla 'distanza generalizzata' di Frei (2014).

Il significato fisico della distanza topografica è legato agli ostacoli incontrati dalle masse d'aria che si muovono orizzontalmente nella bassa atmosfera: a parità di distanza, ci si attende che i punti situati lungo la stessa valle presentino un comportamento simile rispetto ai punti separati da un rilievo topografico. Date due generiche posizioni A e B con altitudini z_A e z_B , la distanza topografica D_t viene calcolata come:

$$D_t = D_c + k_z D_z + k_h D_h$$

$$D_z = |z_B - z_A|$$

$$D_h = \max(\Delta z_i)$$

Dove:

- D_c rappresenta la distanza cartesiana;
- D_z è un termine che dipende dalla differenza di altitudine;
- D_h è la massima differenza di altitudine positiva incontrata lungo il percorso tra i due punti, ovvero il contributo ostacolo effettivo a D_t . D_h viene calcolato discretizzando il percorso tra i punti in intervalli la cui lunghezza orizzontale corrisponde alla dimensione della cella del DEM, e seguendo il percorso dalla quota più bassa a quella più alta.
- Δz_i rappresenta la differenza tra l'altitudine ad ogni passo e quella al punto di partenza. Vengono considerate solo le differenze positive, che rappresentano ostacoli topografici effettivi, mentre le differenze negative (depressioni) vengono escluse. L'esclusione delle conche è presa da Frei (2014), ed evita la sottostima per le stazioni situate, ad esempio, ai lati opposti della stessa valle. L'asimmetria tra ostacoli convessi e concavi è un modo per rappresentare la similarità tra siti collegati attraverso l'atmosfera libera.

- I coefficienti k_z e k_h vengono introdotti come fattori che adattano l'influenza della topografia. Questo tiene conto delle diverse condizioni atmosferiche quotidiane, quando l'anisotropia topografica potrebbe essere più o meno pronunciata.

I coefficienti ottimali da utilizzare vengono stimati mediante una procedura iterativa che minimizza l'errore medio di convalida incrociata.

Il calcolo di D_h viene applicato al profilo altimetrico lungo una linea retta che collega due punti, il che in realtà è un procedimento monodimensionale che quindi comporta una semplificazione della rappresentazione della topografia tridimensionale, portando potenzialmente ad andamenti spaziali irrealistici in output all'interpolazione. Deng e Stull (2005) e Frei (2014) hanno attenuato questa semplificazione e determinato distanze modificate topograficamente dal campo altimetrico completo anziché solo dal suo profilo lungo la linea retta di connessione. Tuttavia, diversi test effettuati sul set di dati utilizzato mostrano che questi andamenti artificiali sono di solito trascurabili. Ciò è dovuto allo schema di interpolazione di Shepard, che include nell'analisi solo un massimo di dieci stazioni vicine alla stazione incognita. Inoltre, l'aggregazione alla griglia finale più grossolana rende ulteriormente trascurabili la maggior parte delle possibili caratteristiche spaziali irrealistiche.

2.5 Eraclito v. 3.0

L'output della procedura è un set di dati in griglia a risoluzione variabile per le precipitazioni giornaliere e le temperature minime e massime negli anni 1961-2010, che copre l'intera regione Emilia-Romagna.

I dati originali sono stati sottoposti a svariate verifiche, al fine di rimuovere dati di bassa qualità, serie temporali non omogenee e correggere eventuali asincronicità. Quest'ultimo test è particolarmente cruciale per un metodo di interpolazione basato sui dati giornalieri ed è stato applicato attraverso un metodo di controllo basato sulla correlazione annuale con una stazione di riferimento. La correzione dei ritardi temporali è rilevante per una rappresentazione giornaliera degli eventi di precipitazione, poiché permette di evitare la frequente comparsa di aree asciutte spurie o aree umide isolate nelle mappe giornaliere.

Sono stati compiuti notevoli sforzi anche nella digitalizzazione dei dati storici e nella definizione di un insieme di stazioni la cui densità fosse ragionevolmente costante nel tempo, un requisito essenziale nell'analisi climatologica.

Il dataset grigliato è stato generato con una nuova procedura di interpolazione, secondo la quale le precipitazioni sono state interpolate tenendo conto delle distanze e degli ostacoli topografici: l'algoritmo Shepard è stato modificato per incorporare gli effetti delle barriere topografiche e delle differenze di elevazione nel calcolo delle distanze tra le posizioni, facendo riferimento, piuttosto che alla distanza cartesiana, alla distanza topografica. La procedura è completamente automatica e l'uso di una griglia triangolata dipendente dall'orografia consente di minimizzare il requisito di spazio di archiviazione, mantenendo contemporaneamente i dettagli spaziali necessari in tutte le zone climatiche.

Nonostante l'uso delle distanze topografiche nell'assegnazione dei pesi alle stazioni migliori leggermente la stima interpolata, gli errori delle precipitazioni dipendono dalla densità delle stazioni. Circa il 90% degli eventi è correttamente rilevato (precipitazione/assenza di precipitazione), con una sottostima media del 10% per eventi medi. Lo smorzamento degli estremi dovuto all'interpolazione e la perdita di dettagli o di piccoli cambiamenti nei modelli di precipitazione, che non vengono catturati o che non sono comunque completamente riprodotti nelle stime interpolate; potrebbero portare a una sottostima di oltre il 50% (in media del 25%) per eventi di precipitazione estivi intensi.

L'archivio di dati creato con questa procedura ha alcune limitazioni inevitabili: ad esempio, i dati raccolti manualmente presentano una qualità inferiore e una maggiore inconsistenza rispetto ai dati raccolti da stazioni elettroniche automatizzate. Sebbene le incongruenze più evidenti siano state rimosse o corrette, alcune asincronicità potrebbero ancora essere presenti, portando all'incertezza nei dati in griglia giornalieri.

Mantenere una densità stabile di stazioni durante un periodo di analisi a lungo termine porta a utilizzare solo una frazione di tutti i dati disponibili in ogni momento. In queste condizioni, le prestazioni giornaliere della metodologia sono ridotte; per questo motivo, il set di dati così ottenuto è adatto per studi climatici, compreso il calcolo degli indici climatici basati su valori giornalieri o per analisi in cui sono necessari valori medi su una scala temporale sufficientemente ampia. Il dataset è meno adatto per analisi di casi di studio giornalieri, per le quali è preferibile l'uso della massima disponibilità di dati in ciascun momento rispetto a una copertura uniforme dei dati nel tempo.

L'applicazione della stessa metodologia alla rete di stazioni in tempo reale più densa e di alta qualità gestita da ARPA-SIMC permette di evidenziare che le prestazioni finali dell'analisi mostrano un netto miglioramento.

2.5.1 Mappe climatiche

Sono stati valutati i valori medi delle variabili climatologiche e i trend temporali, al fine di descrivere il clima regionale e la sua variabilità nello spazio e nel tempo, nel corso dei 50 anni (1961-2010) presi in esame. I risultati sono presentati tramite mappe climatologiche, ricavate interpolando i valori della rete triangolata (TIN) riferiti al baricentro di ciascun triangolo, su una griglia regolare con risoluzione massima equivalente, utilizzata come input del contouring software. Questo può causare una leggera imprecisione nella posizione delle linee di contorno, poiché l'interpolatore IDW si basa solo sulle distanze, senza tener conto delle aree. Gli errori potrebbero accentuarsi dove le dimensioni dei triangoli sono molto variabili (ad esempio tra pianura e colline): i valori dei triangoli più grandi possono essere sottopesati, mentre quelli dei triangoli più piccoli potrebbero essere sovrappesati. Tuttavia, alcuni test preliminari condotti per valutare questo effetto mostrano che gli errori derivanti da questa procedura sono trascurabili. D'altra parte, l'esportazione dei baricentri consente di preservare l'alta risoluzione del set di dati in presenza di orografia complessa.

In Figura 2.6 a e b, sono riportate le mappe climatiche del valore medio delle precipitazioni annuali cumulate e del trend, calcolato utilizzando il test di Mann-Kendall (Mann, 1945; Kendall, 1975) e della sua significatività.

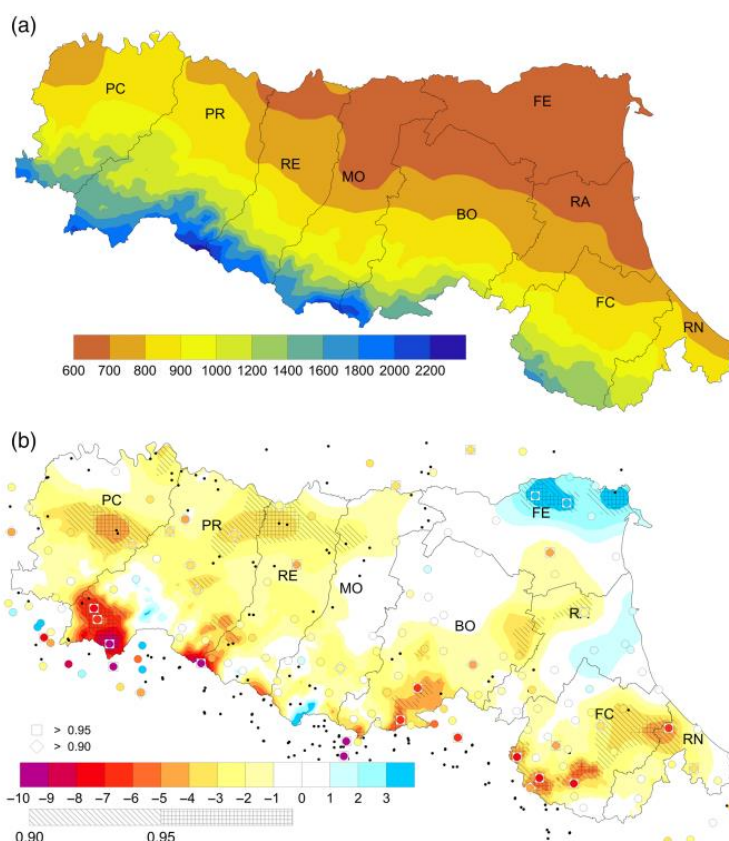


Figura 2.6 (a) Precipitazione media annuale cumulata media nel periodo 1961-2010 (mm), (b) trend della precipitazione media annuale cumulata per il periodo 1961-2010 (mm/anno, punti pieni colorati) e la sua significatività statistica (tratteggio e simboli). I punti neri rappresentano le stazioni utilizzate nell'analisi, con dati insufficienti per il calcolo della tendenza (<80% dei dati nel periodo 1961-2010) (Antolini et al., 2016).

I trend e le loro significatività sono calcolati sia sui dati in griglia che sulle stazioni con dati sufficienti (> 80%); sono rappresentate anche le stazioni con una minore presenza di dati che contribuiscono al set di dati in griglia. In alcuni casi, sono visibili differenze tra i valori di trend sulle stazioni e sulla griglia, a causa della minore disponibilità di dati per le stazioni e del processo di aggregazione.

I valori medi delle precipitazioni annuali cumulate variano da 600 mm, nella pianura orientale, a oltre 2300 mm nelle montagne più alte (tra le province di PR e RE). Le precipitazioni annuali medie sulle pianure non superano i 900 mm, mentre sulle colline e sulle montagne le precipitazioni sono caratterizzate da una variabilità spaziale molto più elevata. È visibile un ampio gradiente sud-ovest/nord-est, principalmente associato all'amplificazione orografica delle precipitazioni durante le transizioni sinottiche; è evidente, infatti, dalla mappa del trend una variabilità spaziale elevata, sia in termini di intensità che di significatività.

In gran parte della regione si osserva una diminuzione generale dei valori medi, sebbene sia statisticamente significativa solo in aree limitate. I trend negativi più intensi si osservano sulle montagne, con valori di quasi -100 mm per decennio sulle cime più alte dell'Appennino. Le

aree con tendenze positive si trovano nella pianura orientale, con significatività solo nella parte settentrionale, con valori di quasi 40 mm per decennio.

I trend osservati e quelli stimati tramite convalida incrociata sono in buon accordo tra loro; i punti nello scatterplot (Figura 2.7) si trovano infatti nello stesso quadrante del piano cartesiano, il che significa che le direzioni dei trend sono, in linea generale, ben rilevate.

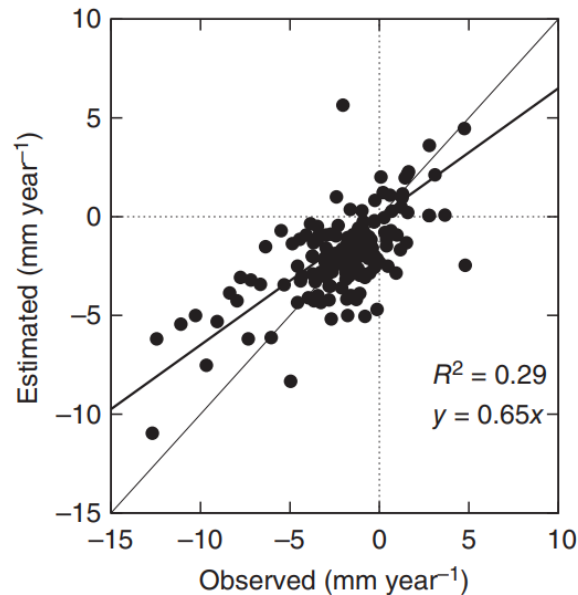


Figura 2.7 Scatterplot dei trend stimati e osservati della precipitazione annuale cumulata (Antolini et al., 2016).

Il valore di R^2 , misura di efficienza, indica che solo il 29% della varianza spaziale del trend viene riprodotta, a causa dell'effetto di smorzamento dell'interpolazione, che influisce anche sui trend. La scorretta riproduzione dei valori di trend osservati potrebbe anche essere dovuta a effetti di ombreggiatura non considerati nel metodo di interpolazione.

2.5.2 Valutazione degli errori

La performance dei metodi precedentemente descritti è stata valutata tramite i seguenti indici di prestazione:

- L'errore medio assoluto (Mean Absolute Error, MAE), la media dei valori dell'errore assoluto:

$$MAE = \frac{1}{n} \sum_{k=1}^n |y_k - o_k|$$

Dove:

- o_k è il k-esimo valore osservato, con $k=1, \dots, n$;

- y_k è il k-esimo valore interpolato;
- n il numero totale di stazioni.

L'errore medio assoluto è una metrica comune che fornisce la quantità media per cui le previsioni differiscono dai dati reali, senza considerare la direzione (positiva o negativa) di tali differenze. Un MAE più basso indica una previsione più accurata, mentre un MAE più alto indica una maggiore discrepanza tra le stime e i dati osservati.

- L'errore relativo composto (Compound Relative Error, CRE), il rapporto medio tra l'errore e la varianza delle misurazioni (deviazione standard al quadrato):

$$CRE = \frac{\sum_{k=1}^n (y_k - o_k)^2}{\sum_{k=1}^n (o_k - \bar{o})^2}$$

Dove \bar{o} è la media dei valori osservati. Questa metrica tiene conto sia dell'entità degli errori che della variabilità dei dati: un CRE più basso indica una previsione più accurata e stabile, mentre un CRE più alto suggerisce una maggiore variabilità nella precisione delle previsioni.

- La percentuale di correttezza (Percent of Correct, PC), un indicatore specificamente progettato per le precipitazioni, che corrisponde alla percentuale di eventi correttamente rilevati in termini di pioggia/non pioggia.

$$PC = \frac{yy + nn}{yy + nn + yn + ny}$$

Dove:

- yy è il numero di eventi di pioggia rilevati correttamente (precipitazioni stimate e osservate sono entrambe positive);
- nn è il numero di eventi di assenza di precipitazione rilevati correttamente (precipitazioni stimate e osservate sono entrambe zero);
- yn è il numero di falsi allarmi (precipitazioni stimate sono positive e precipitazioni osservate sono zero);
- ny è il numero di eventi di pioggia non rilevati (precipitazioni stimate sono zero e precipitazioni osservate sono positive).

L'indice di prestazione PC fornisce una misura della correttezza delle stime ottenute: maggiore è la percentuale di correttezza, maggiore è l'accuratezza del risultato dell'interpolazione.

In particolare, nelle Figura 2.8 sono riportati gli errori relativi dell'analisi nella convalida incrociata, mediati su tutte le stazioni disponibili, raggruppati in base all'intensità e presentati mediante box plot. I due grafici fanno riferimento separatamente all'inverno esteso, quando è più probabile che si verifichino precipitazioni stratiformi, che generalmente apportano pioggia distribuita uniformemente al di sopra di un'ampia zona; e la stagione estiva, quando le precipitazioni sono principalmente convettive, cadono sotto forma di rovesci con cambiamenti rapidi d'intensità e sopra una certa zona per un periodo di tempo limitato.

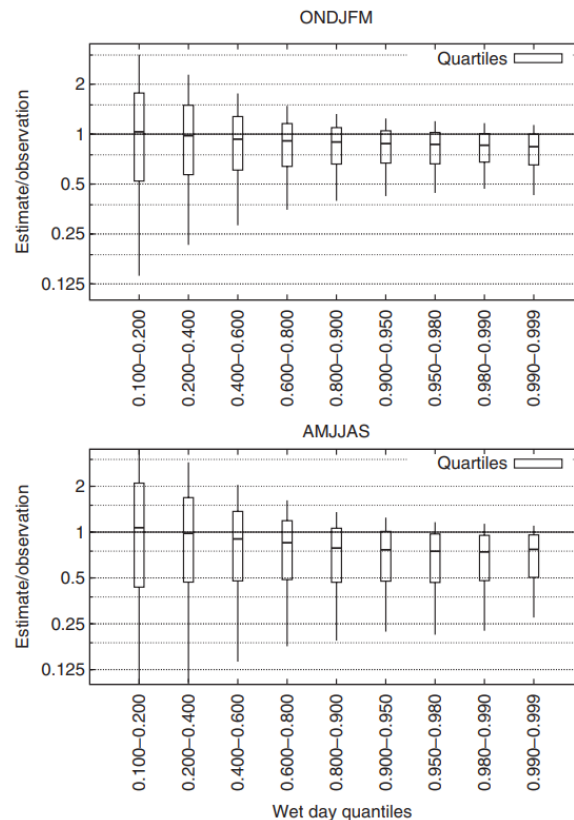


Figura 2.8 Diagrammi a boxplot degli errori relativi per le precipitazioni, calcolati nella convalida incrociata per l'estate e l'inverno (per il periodo 1961-2010) e suddivisi per intensità, utilizzando classi di quantili di giorni piovosi (≥ 1 mm) per ciascuna stazione. Sono rappresentati la mediana, il primo e il terzo quartile e il 10° e 90° percentile (Antolini et al. 2016).

L'errore relativo è il rapporto tra il valore interpolato in una posizione specifica, calcolato nella convalida incrociata, e il valore osservato corrispondente; questo significa che ci si può aspettare di avere grandi errori relativi nelle classi di bassa intensità. Il grafico, infatti, indica che gli errori relativi di solito diminuiscono all'aumentare dell'intensità.

In generale, l'effetto di smorzamento dell'interpolazione comporta una sottostima delle precipitazioni a medio e alto livello di intensità, e una sovrastima a bassa intensità, specialmente in estate. La sottostima media per le precipitazioni intense è del 13% in inverno e del 25% in estate; per le piogge medie è del 7% in inverno e del 10% in estate. Gli errori

sono più elevati in estate che in inverno, per tutte le classi di intensità: la struttura delle piogge convettive estive su piccola scala non è pervenuta nei dati in griglia. Ciò è da imputare al fatto che le prestazioni dell'interpolazione per le precipitazioni dipendono molto dalla densità delle stazioni, densità che è inferiore nella pianura dell'Emilia-Romagna, dove si verificano la maggior parte delle precipitazioni convettive estive (Giuliaci, 1988).

In media, considerando l'intero periodo, circa il 90% degli eventi è correttamente rilevato in termini di presenza/assenza (soglia di 0,2 mm) di precipitazioni. L'utilizzo nello schema di interpolazione della distanza topografica, anziché cartesiana, mostra un piccolo miglioramento per le precipitazioni, indipendentemente dalla stagione, come è possibile osservare nella Figura 2.9, con riferimento agli indicatori di prestazione MAE e PC.

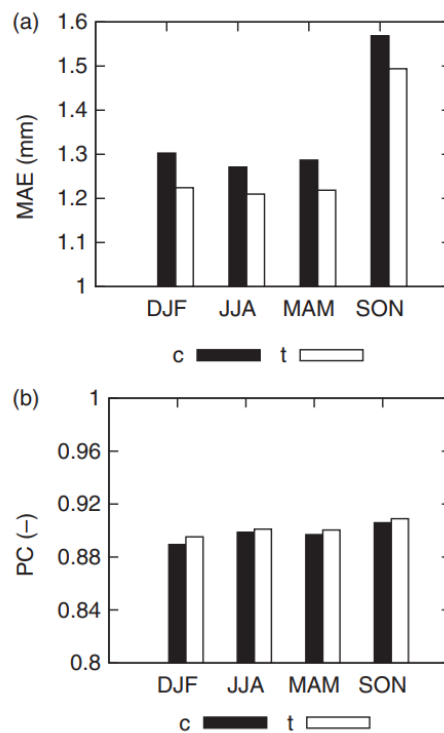


Figura 2.9 Confronto delle prestazioni dell'interpolazione delle precipitazioni utilizzando distanze cartesiane (c) e distanze topografiche (t), in termini di MAE (a) e PC (b) (Antolini et al., 2016).

Capitolo 3: Tecnica sperimentale per la determinazione del fattore di ragguglio areale delle precipitazioni

3.1 Procedura sperimentale di valutazione dell'ARF

L'ARF può essere valutato attraverso una vasta gamma di tecniche. Alcune di queste sono ampiamente applicabili a livello generale, mentre altre sono specifiche per determinate aree geografiche, in cui la stima è strettamente legata alle caratteristiche locali del territorio.

Il presente capitolo si propone di illustrare nel dettaglio la procedura sperimentale adottata per il calcolo del fattore di riduzione delle piogge all'area, con un focus specifico sui bacini imbriferi dell'appennino tosco-emiliano. La metodologia di calcolo è stata messa a punto sulla base degli studi di Penta (1974) e le successive elaborazioni relative al bacino del Fiume Crati ed applicata per un maggior numero di durate di evento piovoso.

Nel presente lavoro, le operazioni applicate per il calcolo del fattore di ragguglio all'area delle precipitazioni (schematizzate nel diagramma di Figura 3.1) prevedono la preliminare elaborazione dei dati provenienti dal dataset ERG5, al fine di poter poi creare gli archivi di riferimento su cui basare la procedura stessa. In Appendice, sono raccolti gli script impiegati per l'elaborazione del dataset ERG5, dei file SHP contenenti i bacini e i sottobacini considerati per il calcolo dell'ARF, e per il calcolo e la rappresentazione dei valori ricavati del fattore di riduzione della pioggia all'area in funzione delle aree prese in esame e delle varie durate selezionate.

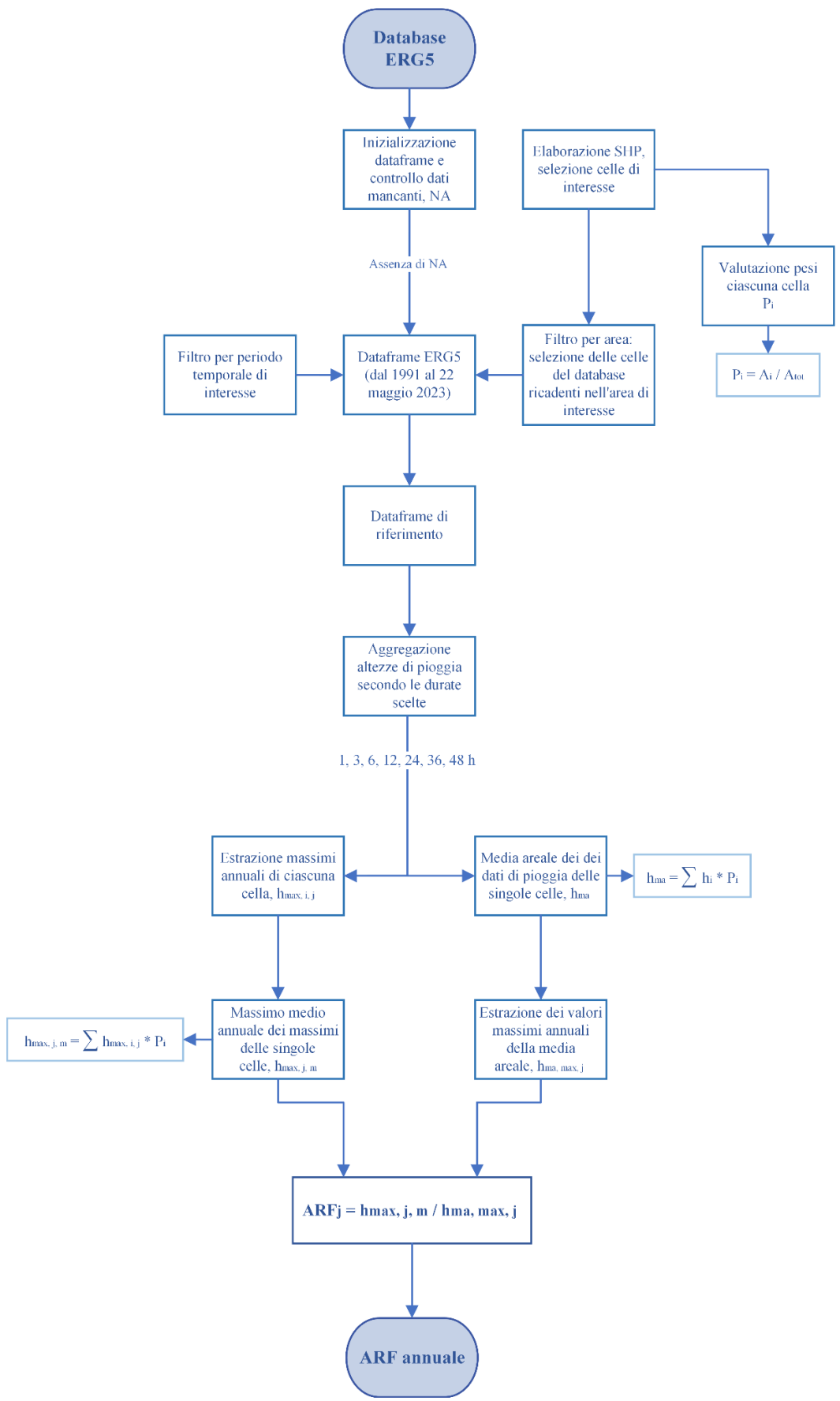


Figura 3.1 Diagramma di flusso del procedimento di valutazione dell'ARF annuale adottato nella Tesi ed implementato attraverso codici di calcolo in linguaggio R.

3.1.1 Dataset base

L'archivio di base utilizzato è costituito da altezze di pioggia mediate su un'area di 25 km² e rilevate ogni ora, per un periodo di riferimento che va dal 1991 al 22 maggio del 2023.

I dati di pioggia sono stati derivati dalla versione 2.0 del dataset meteo orario e giornaliero ERG5, un dataset meteorologico di dati orari e giornalieri per le principali variabili meteorologiche e agrometeorologiche che copre tutto il territorio regionale complessivamente dal 1991 ad oggi.

La raccolta dati è prodotta dall'Osservatorio Clima e l'Unità Territorio e reti di Arpae SIMC ed ottenuta tramite interpolazione spaziale su una griglia regolare a partire dai valori rilevati dalla rete delle stazioni meteorologiche, costituita da 1131 celle di estensione di circa 25 km² che ricoprono l'intera Emilia Romagna. A partire da questa, si è costruito un data frame con il software di programmazione RStudio, contenente i dati di precipitazione a passo orario riferiti a ciascuna cella contenuta nel dataset ERG5 ed al periodo di interesse selezionato.

Nel dettaglio, l'archivio informatico di partenza si costituisce di un numero di file in formato CSV (Comma-Separated Values) pari al numero di celle, file che contengono i valori delle seguenti grandezze orarie interpolate: temperatura dell'aria, precipitazioni, umidità relativa dell'aria, intensità (scalare e vettoriale) e direzione del vento, irradianza solare. Considerando l'altezza di pioggia areale riferita ad un elemento di 25 km² e aggregata con passo orario, i valori disponibili sono espressi in mm e, in particolare, il valore assegnato ad un determinato orario si riferisce alla pioggia caduta nell'ora immediatamente precedente (se in corrispondenza dell'ora 9:00 si legge un valore x, vuol dire che tra le 8 e le 9 in quel punto sono caduti x mm di pioggia).

I file CSV dell'intero dataset sono stati forniti in due cartelle separate, tali per cui la prima cartella contiene i soli dati di altezza di precipitazione accumulatasi a passo orario sulla superficie di una generica cella del dataset dal 1991 al 2000, mentre la seconda cartella contiene dati medi areali riferiti a ciascuna cella e al periodo dal 2001 al 22 maggio 2023 di più grandezze, quali precipitazione, temperatura, irradiazione, intensità del vento. In ulteriore dettaglio, questi si costituiscono di quattro colonne:

- La colonna DATE, che contiene il valore specifico della data, anno-mese-giorno;
- La colonna HOUR, che contiene l'indicazione oraria nell'arco giornaliero in formato numerico, da 0 a 23;

- La colonna VARIABLE, che contiene l'indicazione della grandezza media riportata, che nel primo caso è la sola altezza di pioggia PREC, mentre nel secondo assume valori quali PREC, appunto precipitazione, TEMP, temperatura, RH, irradiazione, e WIND, vento;
- La colonna VALUE, dove sono riportati i valori numerici relativi alla grandezza considerata.

La realizzazione dell'archivio di base è stata articolata in modo includere il controllo, a favore di sicurezza, della possibile presenza di dati mancanti o non disponibili, NA. Si potrebbe infatti verificare l'eventualità che un pluviometro non funzioni o che il suo funzionamento sia incompleto, per cui in corrispondenza di essa l'archivio dei dati originali presenterebbe lacune. In tal caso sarebbe necessario prendere opportune decisioni per il loro trattamento, come la ricostruzione tramite, ad esempio, tecniche di multiregressione lineare; o l'eliminazione della riga ad essi corrispondente. D'altra parte, i valori delle variabili meteorologiche raccolti nel database ERG5, utilizzato come fonte dati, sono sottoposti a scrupolosi controlli da parte della Struttura IdroMeteoClima Arpa, per cui il risultato atteso di questa verifica è l'assenza di dati mancanti.

In dettaglio, si è proceduto con l'inizializzazione di un data frame costituito da una colonna contenente il riferimento temporale a passo orario (data e ora espressi nel fuso orario Coordinated Universal Time (UTC)), e da un numero di colonne pari al numero di celle del grigliato ERG5, i cui elementi sono valori mancanti, NA. In tali colonne sono state archiviate opportunamente le altezze di precipitazione a passo orario contenute nei CSV associati a ciascuna cella. In dettaglio, ciascuna di queste colonne siano riportati integralmente i dati di precipitazione relativi una cella specifica nel grigliato. Infine, è stato applicato il controllo per l'individuazione di eventuali dati mancanti, che, come previsto, ha dato esito negativo.

Si è quindi ottenuto in RStudio un data frame che racchiude integralmente le altezze di pioggia a passo orario per tutte le celle di cui si costituisce il database ERG5, per un periodo che va dal 1991 fino al 22 maggio 2023. Il data frame è correlato di riferimento temporale a passo orario e di opportuna indicazione della cella specifica cui fanno riferimento i valori di ciascuna colonna, tramite l'impiego dei codici identificativi con cui sono nominate le celle nel database ERG5 originale.

3.1.2 Dataset di interesse

Una volta creato il data frame complessivo, corrispondente dataset ERG5 v 0.2 con riferimento ai dati di precipitazione a passo orario, si procede con la selezione dei dati di interesse in funzione dell'area e del periodo in riferimento dei quali si vuole eseguire il calcolo dell'ARF.

Al fine di selezionare i valori di altezza di precipitazione cumulata in 1 ora, relative alle celle che ricadono nell'area di interesse, si è proceduto con l'elaborazione di file SHP contenenti l'area di interesse. In R, questa si realizza utilizzando il pacchetto “*terra*”, specializzato nella manipolazione, visualizzazione e analisi dei dati geospaziali, che fornisce funzioni e strumenti per lavorare con raster, vettori e dati spaziali, inclusi operazioni di estrazione, aggregazione, calcolo di statistiche, eccetera.

Nel dettaglio, una volta selezionato lo SHP poligonale dell'area di riferimento, è stato possibile ritagliare il layer vettoriale del grigliato ERG5 (Figura 3.2) rispetto ad esso.

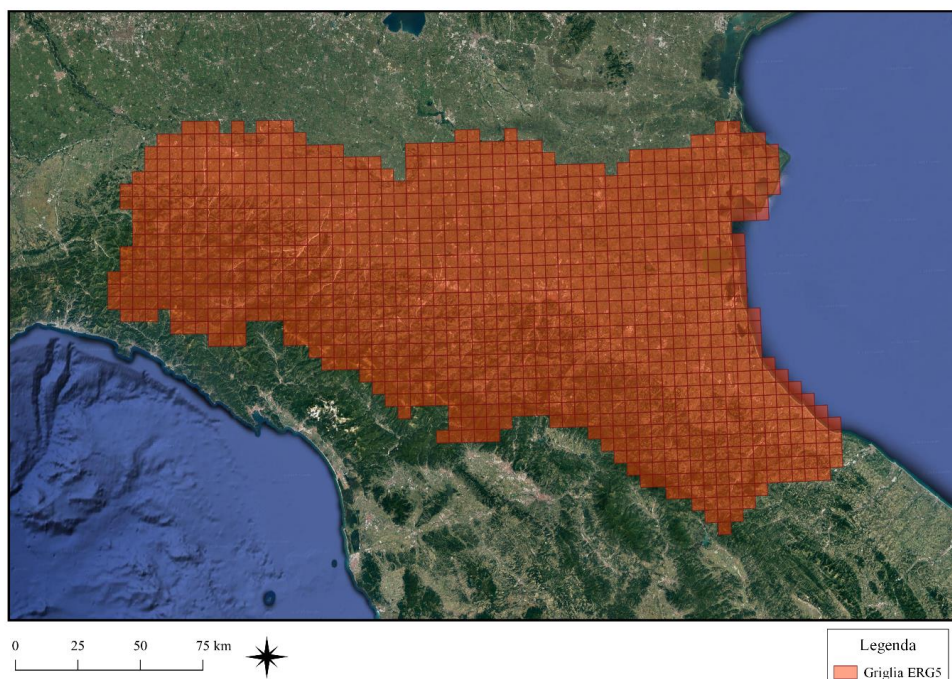


Figura 3.2 Layer vettoriale multipoligonale del grigliato ERG5 in QGIS.

In sintesi, viene identificata la porzione del grigliato, e quindi le celle, le cui informazioni identificative sono riportate nel file SHP del grigliato, che si sovrappone al poligono specificato. Si ottiene un vettore multipoligonale, i cui poligoni sono le frazioni delle celle ricadenti nell'area in esame, per il quale viene calcolata l'estensione in km^2 . Tramite quest'ultima, sono stati valutati i pesi, P_i , assegnati alle celle evidenziate, operazione descritta nel dettaglio nei paragrafi 3.3 e 3.4 del corrente capitolo.

Sulla base delle celle individuate si applica un primo filtro al data frame per la selezione delle colonne contenenti i dati di precipitazione a passo orario ad esse relativi.

Segue una seconda operazione di filtrazione relativa al periodo di riferimento di trattamento dei dati, pari a 32 anni di osservazioni: dal 1991 al 2022.

3.1.3 Aggregazione delle altezze di precipitazione

I dati selezionati sono altezze di pioggia misurate con passo orario e mediate su aree di 25 km²; le durate d prese in considerazione per il calcolo dell'ARF annuale sono sette (1, 3, 6, 12, 24, 36, 48 ore): si procede all'aggregazione dei valori con riferimento a queste durate.

L'operazione di aggregazione è schematizzabile come la semplice somma di valori in successione temporale. Poiché l'espressione "precipitazione di durata di n ore" sottintende un valore cumulato della pioggia caduta in n ore consecutive all'interno di un evento protrattosi per un periodo di lunghezza anche differente, quando si procede all'aggregazione ad una certa durata d si tiene conto non solo delle piogge durate d , ma anche di quelle di durata inferiore e superiore. Dunque, l'operazione di somma in successione temporale prescinde dalla durata dell'evento e tiene conto, a seconda che si abbia a che fare con eventi più brevi o più lunghi, delle altezze di pioggia totali o relative al sottoperiodo di lunghezza pari a d .

Se ne ottengono sette dataframe, contenenti i dati di precipitazione cumulata, ciascuno secondo una delle durate considerate, $h_i(d)$.

3.1.4 Media dei massimi annuali

Per ciascun data frame ottenuto, si estraggono i valori massimi annuali di ciascuna cella i -esima in ogni anno j -esimo considerato, $h_{max, i, j}(d)$.

I data frame dei massimi annuali per ciascuna durata scelta vengono ulteriormente elaborati, in modo tale da calcolare, a partire da essi, la media pesata (con pesi P_i ottenuti come rapporto tra l'area della i -esima cella che ricade nell'area di interesse, e l'estensione totale dell'area di interesse) dei valori massimi relativi alle singole celle, ottenendo quindi un unico valore medio del massimo annuale di precipitazione, $h_{max, j, m}(d)$:

$$h_{max, j, m}(d) = \sum_{i=1}^n P_i * h_{max, i}(d)$$

Con n numero totale di celle che ricadono nell'area di interesse.

3.1.5 Estrazione della precipitazione media areale massima annuale

Il punto di partenza, in questo caso, è il data frame delle serie orarie di precipitazione delle singole celle, filtrato sulla base del periodo e dell'area di interesse. La precipitazione media areale a passo orario al generico istante di tempo t, viene valutata come la media pesata, con pesi P_i , delle altezze di pioggia a passo orario relative alle celle selezionate e all'istante di tempo t:

$$h_A(d = 1h) = \sum_{i=1}^n P_i h_i(d = 1h)$$

Una volta ricavati i valori di precipitazione media areale a passo orario in tutto il periodo di riferimento scelto si passa all'aggregazione dei dati per le sette durate prese in esame, in maniera analoga a quanto fatto nelle operazioni precedenti con riferimento alle singole celle. Si ottengono nuovamente sette data frame che raccolgono i valori di precipitazione media areale cumulata secondo ogni durata considerata in sede di elaborazione.

Anche per i dati areali si estraggono gli eventi che, fissati anno j e durata d, presentano le maggiori altezze di precipitazione, $h_{A, \max, j}(d)$.

3.2 Il calcolo dell'ARF

Gli archivi ricavati, per ogni durata considerata, di valore medio annuale dei massimi delle singole celle e dei massimi annuali delle precipitazioni medie areali, forniscono la base per la determinazione del fattore di riduzione all'area delle precipitazioni medio annuale, con riferimento all'area selezionata e per ogni durata considerata.

Solitamente ciò che è incognito è il massimo annuale delle precipitazioni medie areali per assegnata durata, e il coefficiente annuale ARF_j viene utilizzato per una sua stima attraverso l'espressione seguente:

$$h_{A, \max, j}(d) = ARF_j(A, d) \cdot h_{\max, j, m}(d)$$

Il massimo annuale delle precipitazioni medie areali è il prodotto tra ARF_j e la media pesata dei valori massimi annuali per le singole celle.

A partire da questa relazione, nel presente lavoro di elaborazione, si è proceduto considerando che l'incognita sia ARF_j e quindi il risultato finale non è la determinazione del valore medio di precipitazione areale, ma il fattore di riduzione annuale stesso.

$$ARF_j(A, d) = \frac{h_{A,max,j}(d)}{h_{max,j,m}(d)}$$

Mediante tale procedura, è stato possibile ottenere una stima con valenza annuale del fattore di riduzione della pioggia all'area, con riferimento ad un periodo di 32 anni di osservazioni, dal 1991 al 2022.

3.3 Determinazione della pioggia media puntuale su un'area assegnata

Data un'area A , su cui si trovano n pluviometri, l'altezza di precipitazione media areale, ovvero la precipitazione media caduta sull'intera area A contribuente al deflusso, ha la seguente espressione:

$$h_A = \sum_{i=1}^n w_i h_{p,i}$$

Dove $h_{p,i}$ è l'altezza di pioggia puntuale di durata 1 ora e w_i è il peso dell' i -esimo pluviometro.

L'elaborazione congiunta delle misure puntuali è di grande importanza per una stima accurata e attendibile della precipitazione media areale, ovvero della precipitazione media caduta su un'intera area A . Inoltre, la precisione legata alla sua stima si ripercuote sul risultato di eventuali valutazioni che richiedano il suo impiego come input, come ad esempio la determinazione della portata di piena come output di un modello afflussi-deflussi, che trasforma le piogge in portate e che ha in input, per l'appunto, la pioggia media areale. L'approccio più semplice per la stima della precipitazione media areale, h_A , consiste nel calcolo del valore della media aritmetica dei valori disponibili. In questo caso, i ponderatori sono tutti uguali, e pari ad $1/n$, se n è il numero dei dati disponibili.

$$h_A = \frac{\sum_{i=1}^n h_{p,i}}{n}$$

Tuttavia, questo metodo assume che ciascun dato sia ugualmente rappresentativo del valore areale. Questa assunzione può rivelarsi giustificata e porta a risultati soddisfacenti solo in rari casi in cui si ha una densità di pluviometri sufficientemente elevata; si dimostra inaccurata nel

caso in cui alcune stazioni siano periferiche, oppure poste all'esterno del bacino, rispetto ad altre, collocate in posizione più centrale.

Al fine di ovviare ai limiti più evidenti del metodo basato sulla media aritmetica, è stato introdotto un secondo metodo, di gran lunga più utilizzato per fini idrologici, che si basa sui poligoni di Thiessen, o topoietai. Il metodo dei poligoni di Thiessen è una tecnica tra le più classiche per spazializzare un dato climatico, in questo caso il dato di pioggia, ovvero per ottenere la valutazione di una precipitazione riferita ad un'area, a partire da informazioni che sono puntuali, poiché si riferiscono al punto dove è collocato il pluviometro.

Questo conduce al calcolo della precipitazione media areale come una media pesata, sotto l'ipotesi la misura di ogni strumento sia rappresentativa per un'area che si estende radialmente dal pluviometro stesso fino alla semidistanza dallo strumento adiacente, in ogni direzione. In altre parole, la regione di studio A è suddivisa in n zone di influenza, A_i , una per ciascuna stazione di misura. La zona di influenza individua la regione dei punti più vicini allo strumento i-esimo che a qualsiasi altra strumento. Queste regioni ad ugual pioggia possono essere individuate tramite un metodo geometrico, disegnando bisezioni perpendicolari alle linee di giunzione di stazioni adiacenti, e quindi trovando la rete di sensori, che intersecandosi compongono un numero di poligoni pari al numero degli strumenti, sottesi dai poligoni stessi. Questa operazione è ottimizzata da strumenti informatici, come gli argomenti disponibili in ambiente GIS.

Il ponderatore P_i relativo alla stazione A_i viene infine calcolato tramite il rapporto fra l'area A_i e l'area complessiva di riferimento A:

$$P_i = \frac{A_i}{A}$$

Con P_i numero adimensionale compreso tra 0 e 1.

Il metodo di Thiessen prende in esame anche i pluviometri esterni all'area in esame, con l'accorgimento di considerare, per i poligoni di confine, l'area di competenza A_i come l'intersezione della superficie del poligono i-esimo con l'area di interesse A. L'altezza di precipitazione media areale si ricava come:

$$h_A = \frac{\sum_{i=1}^n A_i h_i}{A} = \sum_{i=1}^n P_i h_i$$

Il metodo dei topoietai è un metodo accurato, che però non è flessibile a variazioni di configurazione della rete pluviometrica, eventualità che non risulta realizzarsi raramente, basti pensare alla frequenza del mancato funzionamento di uno strumento. Non è applicata, inoltre,

alcuna distinzione in funzione delle condizioni orografiche dell'area di riferimento, per cui zone montuose e zone pianeggianti sono trattate ugualmente. Infine, non è permessa la continuità della distribuzione di pioggia, proprio per le linee che separano i poligoni, a causa delle quali si ha un brusco passaggio da un valore di precipitazione a un altro.

Un metodo ulteriore, per la valutazione della pioggia media areale, è il metodo delle isoiete, che ovvia, in parte, agli svantaggi del metodo dei poligoni di Thiessen. In pratica, vengono tracciate in pianta le linee di uguale altezza di precipitazione, sulla base dei valori registrati dalle stazioni ed eventualmente interpolati. A partire da queste linee, si procede con la misura planimetrica dell'area individuata da due isoiete successive, A_i , e si suppone che su di essa sia caduta una precipitazione uniforme nello spazio, pari alla semisomma dei valori tipici delle isoiete estreme, h_i .

L'espressione dell'altezza di pioggia media areale è quindi analoga a quella che si utilizza nell'ambito del metodo dei poligoni di Thiessen:

$$h_A = \frac{\sum_{i=1}^n A_i h_i}{A} = \sum_{i=1}^n P_i h_i$$

Dove però la precipitazione h_i ha significato diverso, ovvero la precipitazione assegnata alla i -esima area individuata da due isoiete consecutive.

Il metodo delle isoiete è un metodo flessibile, a differenza di quello dei topiети. D'altra parte, la sua attendibilità è funzione della densità della rete pluviometrica: i suoi risultati sono affidabili nel caso di reti pluviometriche molto raffittite; inoltre è un metodo complesso, poiché la procedura va ripetuta per ogni altezza di pioggia e la mole di lavoro è notevole, anche se ridotta dal supporto di mezzi informatici.

La precipitazione puntuale può essere mediata all'area secondo altre metodologie, come la tecnica della distanza reciproca (Wei e McGuinness, 1973), secondo la quale si assume che il valore di precipitazione rilevato in un punto A abbia un'influenza su un punto B privo di misura che sia proporzionale al reciproco della distanza planimetrica tra i due punti.

In conclusione, le tecniche di valutazione dei pesi per il calcolo della precipitazione media areale sono svariate; a tal merito, assume rilievo uno studio di Sing e Chowdury, che dal confronto dei vari metodi disponibili in letteratura per la valutazione della pioggia media areale, osservarono che questi forniscono risultati tra loro compatibili, se applicati a precipitazioni puntuali di durata elevata.

3.4 Dati di precipitazione a passo orario ERG5

Le altezze di pioggia del database ERG5 non sono misurazioni puntuali, bensì il risultato dell'interpolazione delle misurazioni puntuali di precipitazione a passo orario puntuali (Capitolo 2), previo opportuno controllo, su griglia regolare 5x5 km.

In altre parole, il dato originale stesso è una precipitazione media areale a passo orario su un'area di 25 km². È a partire da questi valori che si ottiene la stima della precipitazione, per una data durata dell'evento di pioggia, che interessa un'assegnata area A. Nella loro trattazione, si è introdotta la semplificazione di poter considerare la i-esima cella ERG5 come unità elementare, per cui questi dati medi areali sono stati considerati come dati puntuali, in quanto, sebbene questi corrispondano ad una riduzione del valore puntuale, ottenuta tramite interpolazione, un'area di 25 km² è di gran lunga inferiore alle aree complessive dei bacini considerati.

Per quanto concerne l'assegnazione dei pesi, preso a riferimento il metodo dei topoi, è stato considerato per il loro calcolo che la valenza della misura di precipitazione disponibile sia estendibile alla cella cui essa è relativa e quindi all'estensione di quest'ultima. Ne consegue la valutazione del peso P_i della i-esima cella secondo la:

$$P_i = \frac{A_i}{A}$$

Il peso P_i è valutato come il rapporto tra l'area della cella inferente all'area in esame, A_i , e l'area oggetto di studio stessa, A.

Capitolo 4: Risultati: il fattore di riduzione areale ARF(A, d) sui bacini esaminati

4.1 Risultati sperimentali

La procedura sperimentale adottata è stata applicata per la stima del fattore di riduzione areale delle precipitazioni, considerando durate di 1, 3, 6, 12, 24, 36, 48 ore e ai bacini e sottobacini dell'Appennino tosco-emiliano dei seguenti corsi d'acqua: F. Reno (fino a circa 3500 km²), F. Taro (intorno ai 1400 km²), F. Trebbia (fino a circa 920 km²), F. Nure (350 km²). Ne consegue la stima dell'ARF per aree che vanno da 25 a 3500 km² per un periodo di studio di 32 anni, dal 1991 al 2022. Da notare che il limite inferiore imposto all'area coincide con l'area delle celle elementari del database ERG5: non avrebbe senso infatti considerarle nelle operazioni di stima, in quanto le altezze di precipitazione, presa come puntuale, ad esse corrispondenti è in realtà una precipitazione ragguagliata sui 25 km² delle celle stesse.

Questa sezione si propone di illustrare i risultati in termini di ARF annuale ottenuti con riferimento al bacino del Reno e relativi sottobacini, in numero pari a 20.

Il bacino preso a riferimento corrisponde principalmente al bacino montano del Reno; sono stati esclusi dall'analisi gli affluenti romagnoli, come evidenziato in Figura 4.1, dove si osserva il layer vettoriale preso in esame.

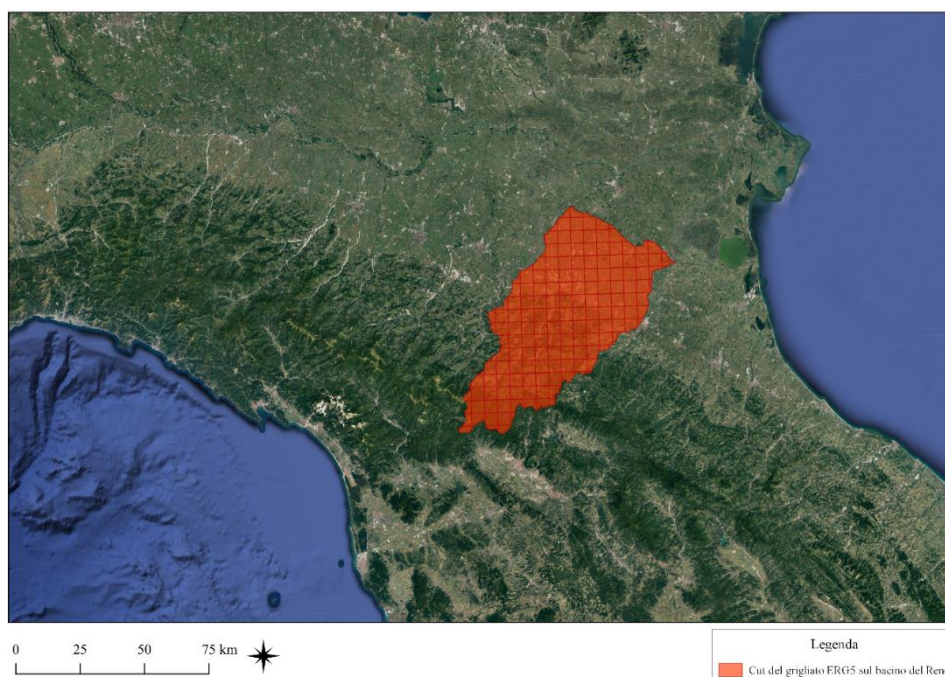


Figura 4.1 Visualizzazione in QGIS della porzione del grigliato ERG5 corrispondente al bacino del Reno (esclusi gli affluenti romagnoli).

Per completezza di illustrazione, nella Figura 4.2 sono evidenziati tutti i bacini principali considerati.

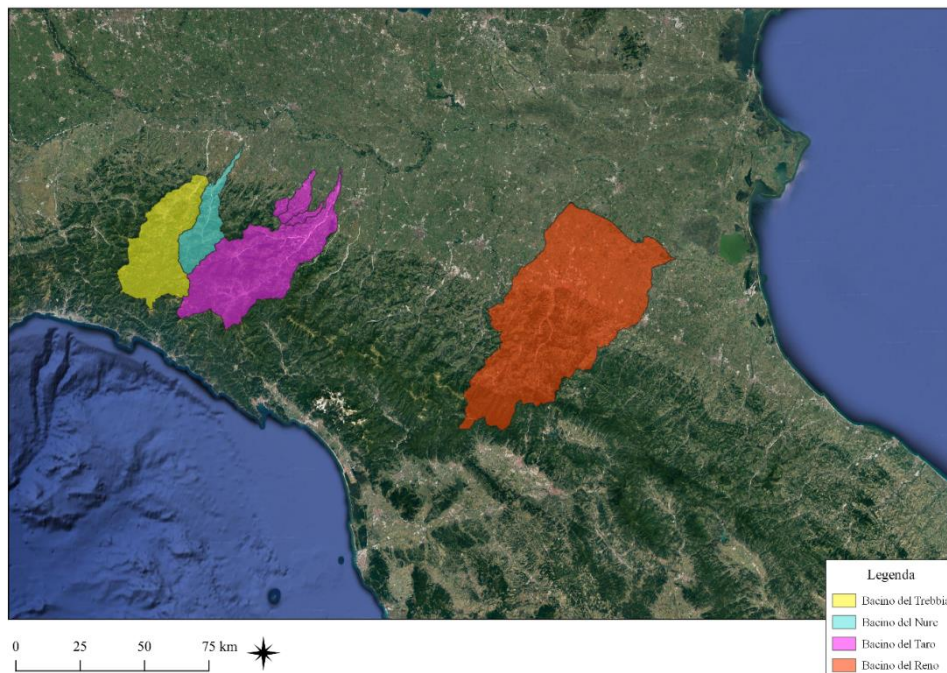


Figura 4.2 Visualizzazione in QGis dei bacini appenninici principali presi in esame, Nure, Trebbia, Taro e Reno.

4.1.1 Risultati sperimentali: $ARF(A, d)$ medio annuale

Per ciascuna durata considerata, sono stati rappresentati gli scatterplot del valore medio della distribuzione costituita dai 32 risultati sperimentali, ARF annuale, per una durata fissata, in funzione dell'area (Figura 4.3).

$$ARF(A, d) = \frac{\sum_{j=1}^{32} ARF_j(A, d)}{32}$$

È evidente che per le brevi durate, i valori sperimentali tendono a decadere più rapidamente con l'area rispetto che alle lunghe durate, comportamento, in generale, imputabile ai diversi tipi di eventi di pioggia che si verificano in diverse stagioni, e quindi in questo caso alla natura prevalentemente convettiva e della limitata estensione spaziale degli eventi di breve durata. Gli eventi convettivi, solitamente caratterizzati da durate inferiori a un'ora, tendono a essere più localizzati e intensi, mentre gli eventi frontali su vasta scala, con durate più lunghe, si caratterizzano per una distribuzione più uniforme della pioggia su un'area più estesa.

Di conseguenza, si osserva che il fenomeno di abbattimento areale della precipitazione è consistente soprattutto per precipitazioni di breve durata su vaste aree; mentre per le lunghe

durate, l'ARF tende a 1, con un decadimento con il crescere dell'area meno pronunciato: gli eventi di pioggia con lunghe durate sono appunto riconducibili ad eventi frontali su larga scala, che di solito si verificano in inverno, e non diminuiscono molto in intensità con l'incremento dell'area.

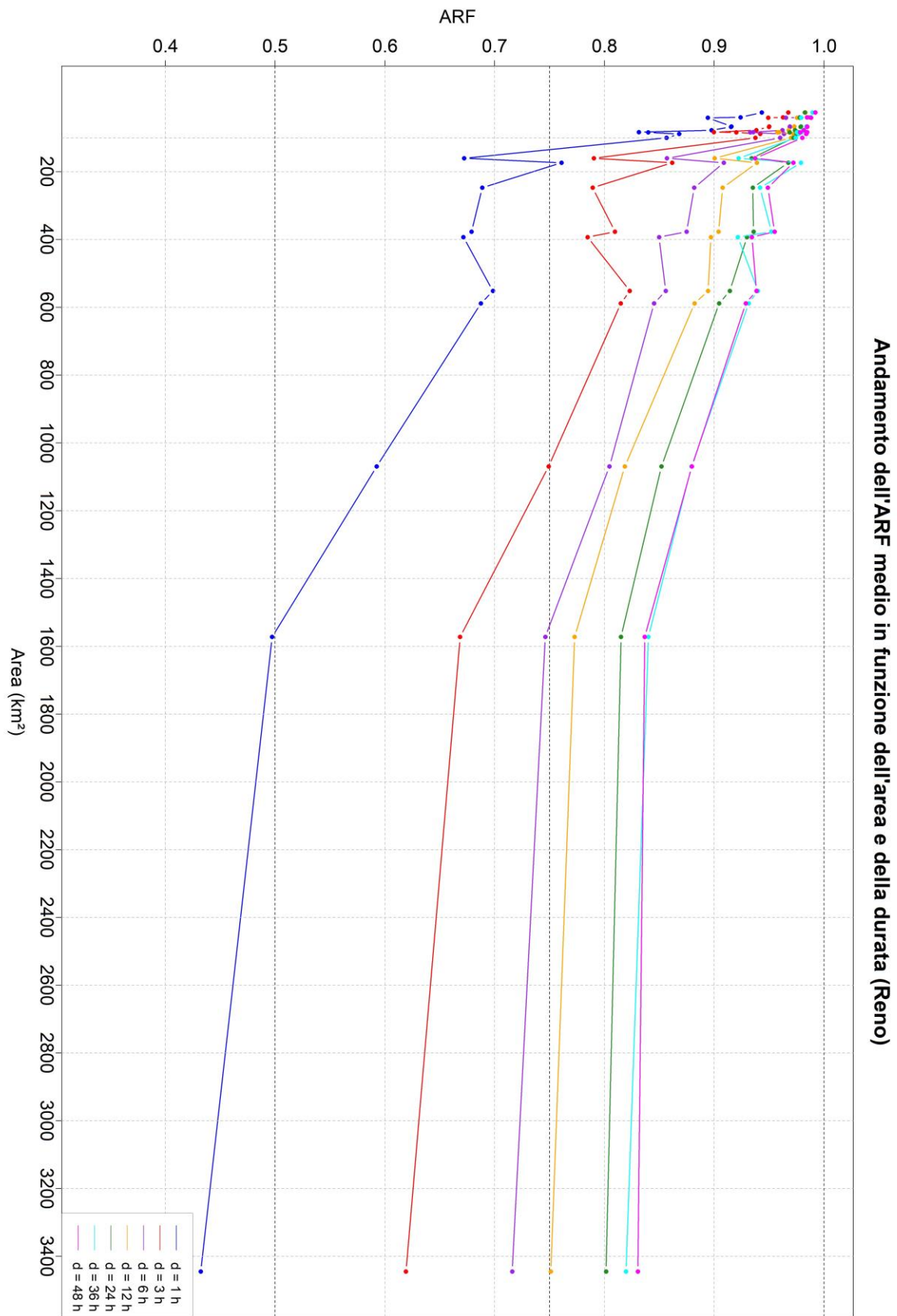


Figura 4.3 Valori sperimentali dell'ARF medio nei 32 anni considerati, al variare della durata, per il bacino del Reno e relativi sottobacini.

In Appendice sono riportati i grafici dell'ARF medio annuale per ciascuna durata per i restanti casi esaminati.

4.1.2 Boxplot della distribuzione dell'ARF(A, d) medio annuale

Per evidenziare la variabilità complessiva dei valori annuali dell'ARF, ed eventualmente identificare valori anomali, sono stati realizzati dei grafici diagnostici appropriati che riassumono visivamente le caratteristiche e il comportamento delle serie stesse, ovvero i boxplot.

Questo tipo di grafico fornisce una visualizzazione dei dati che riassume la distribuzione di un campione utilizzando indici di posizione. La rappresentazione si basa su un rettangolo limitato dai percentili 25% e 75% (il 1° e il 3° quartile), con una divisione interna in corrispondenza della mediana, nota anche come percentile 50% (il 2° quartile). Dal rettangolo si estendono due linee, chiamate "baffi" che raggiungono il valore minimo e massimo del campione.

Nelle Figure 4.4 a-g, sono riportati i boxplot per ciascuna durata, con riferimento al Reno e ai suoi sottobacini.

Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 1 h (Reno)

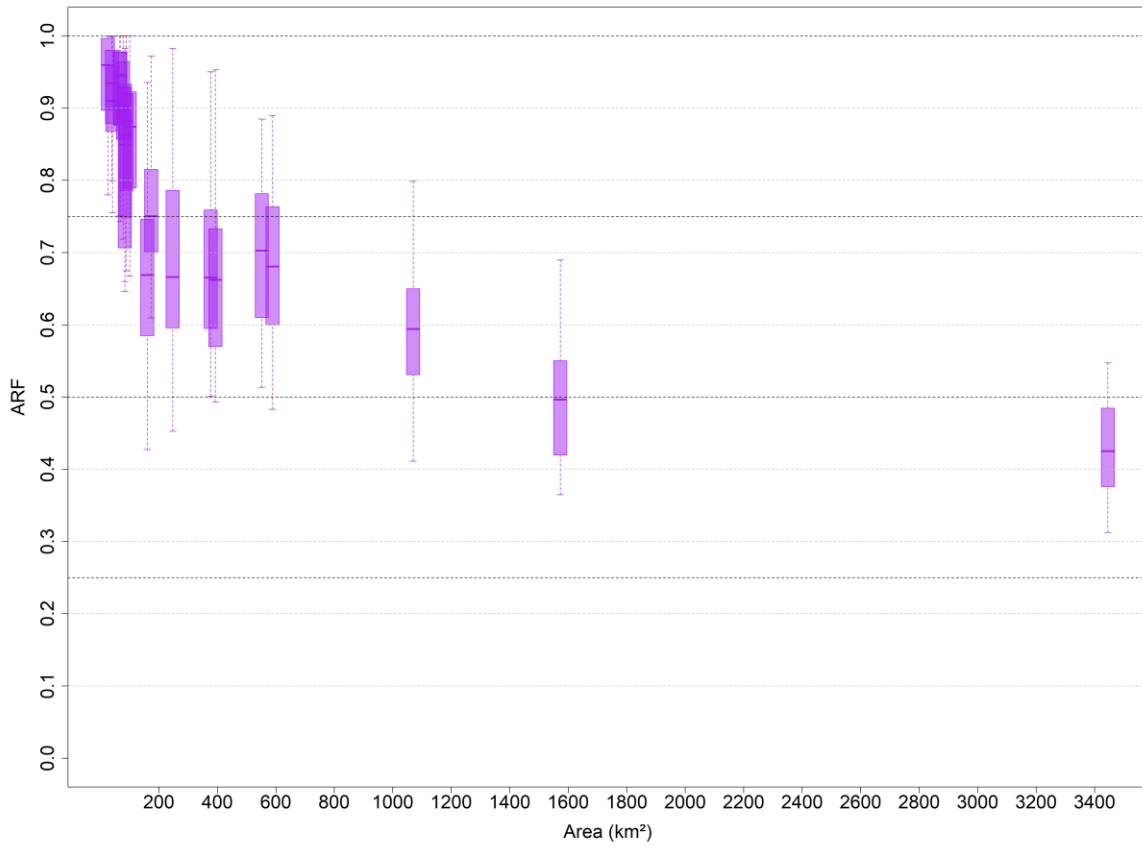


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 3 h (Reno)

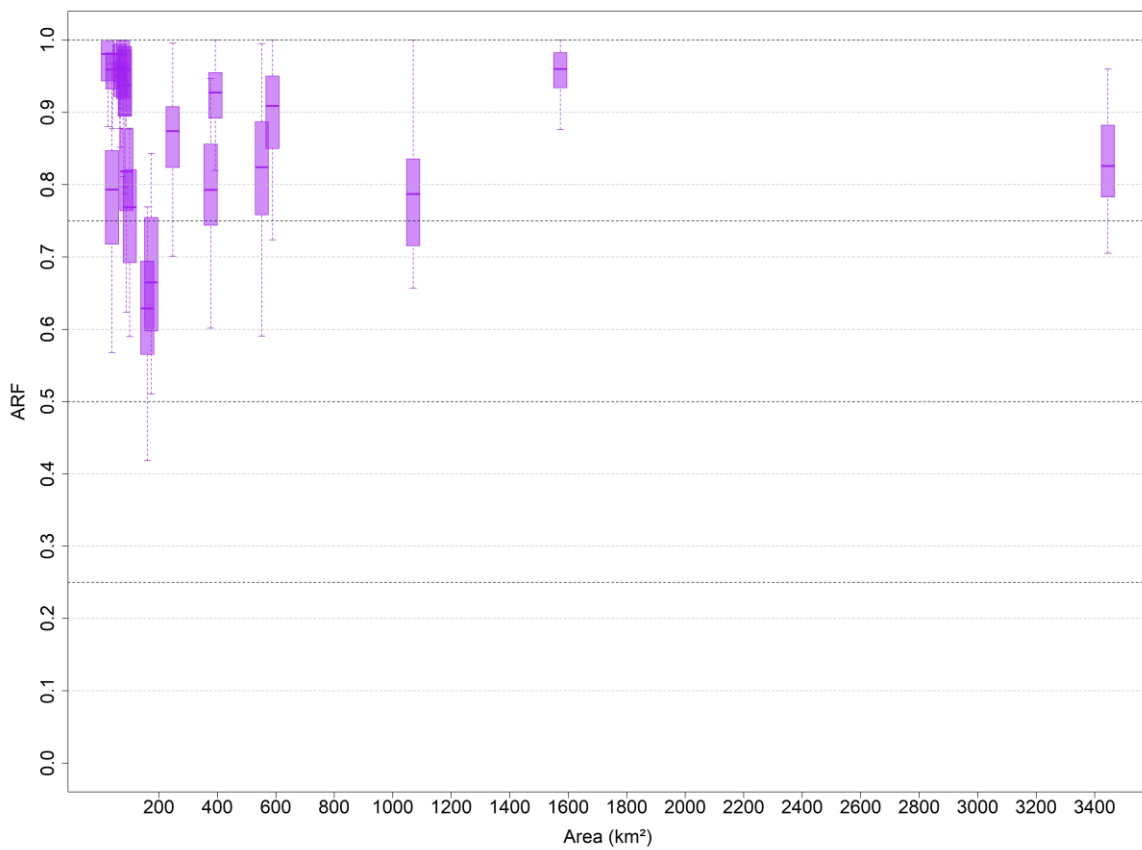


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 6 h (Reno)

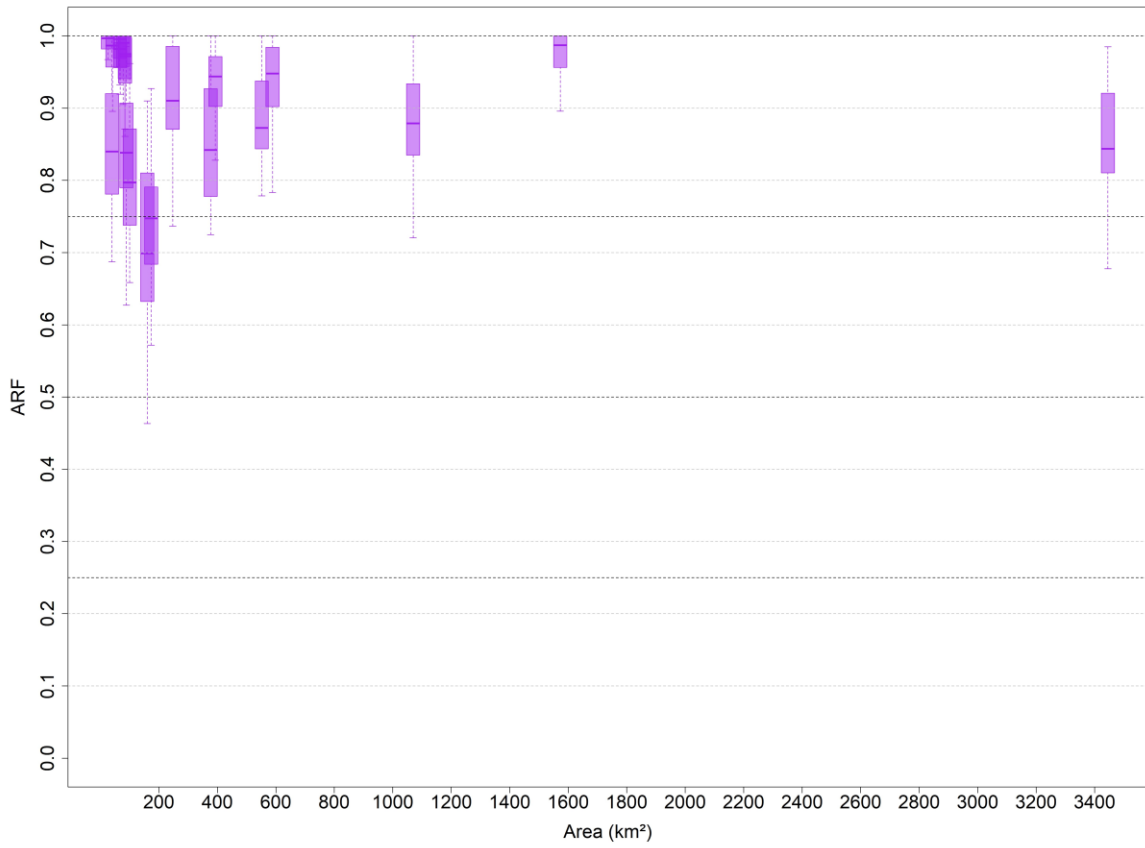


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 12 h (Reno)

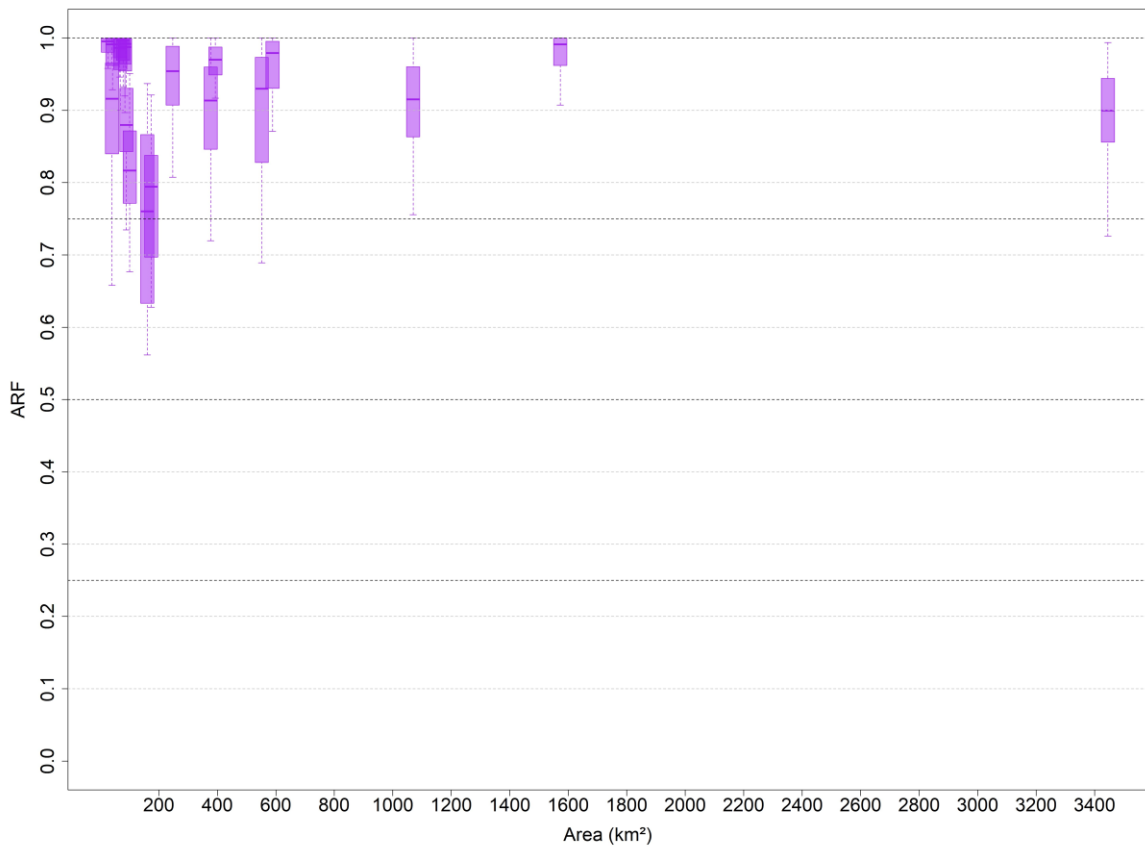


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 24 h (Reno)

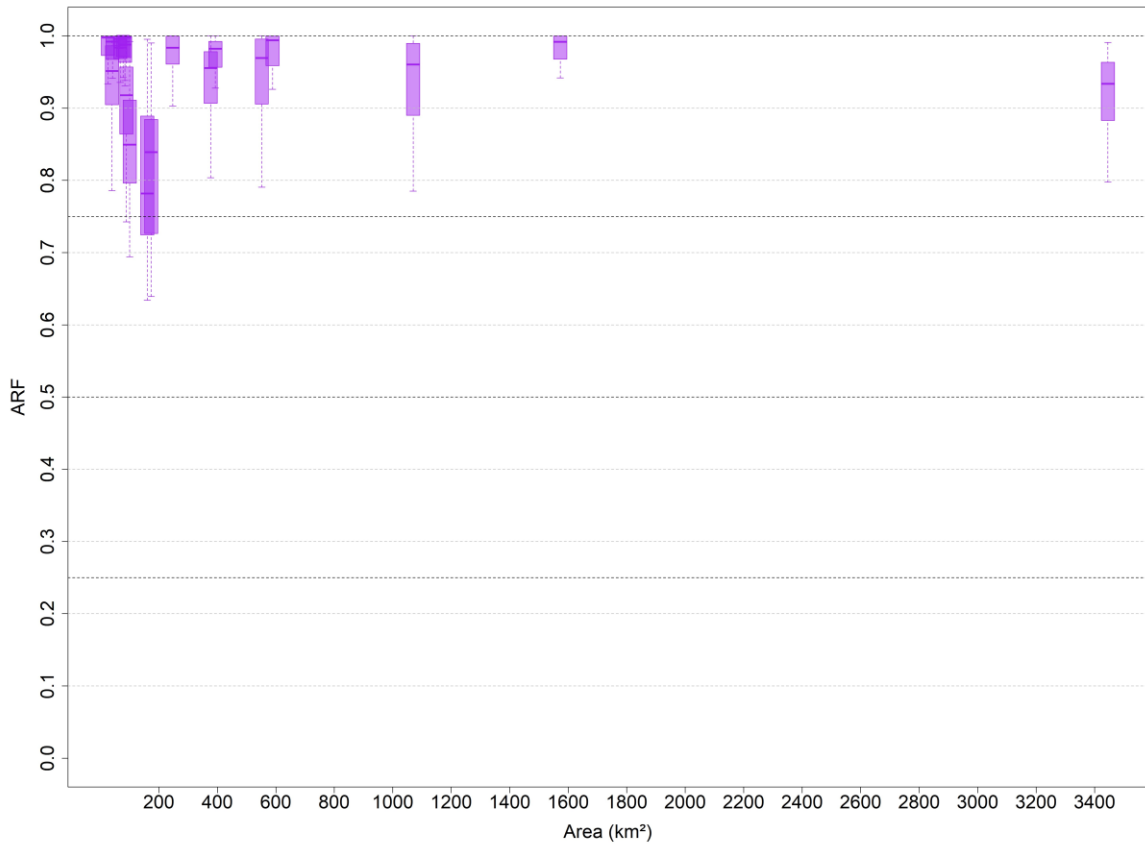


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 36 h (Reno)

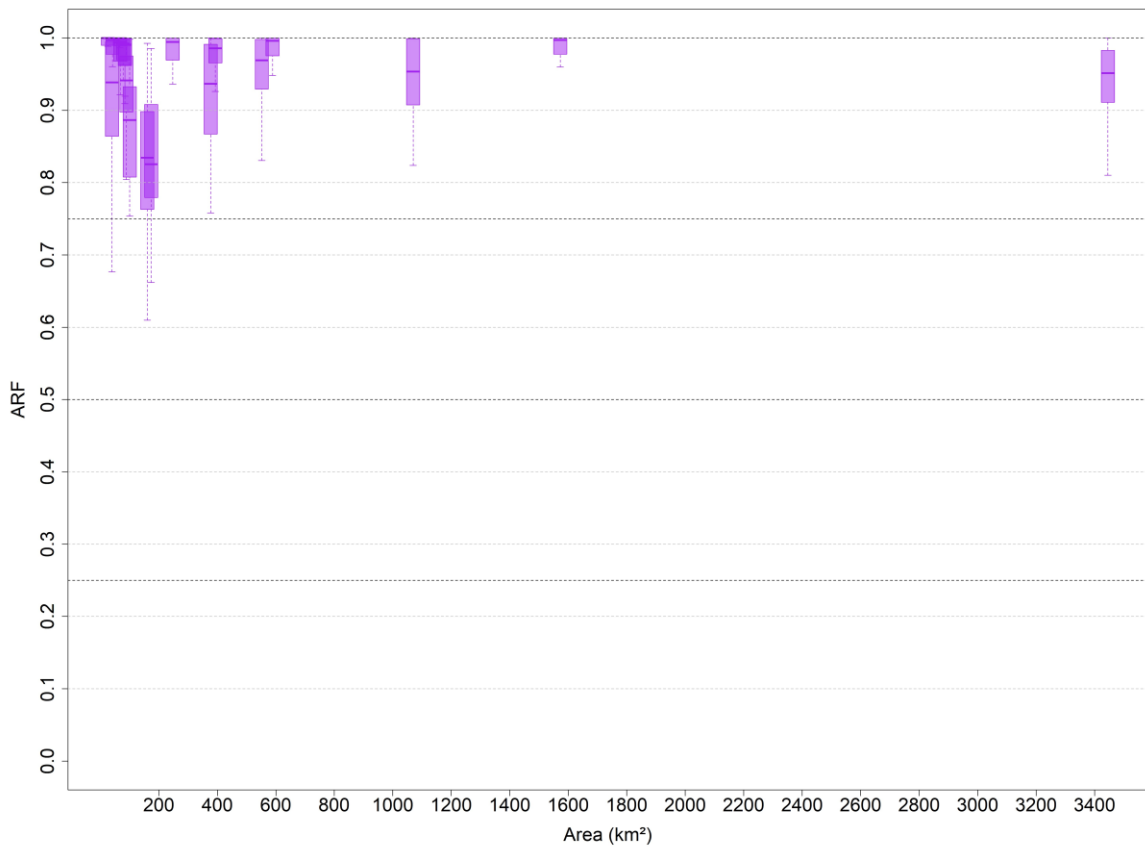


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 48 h (Reno)

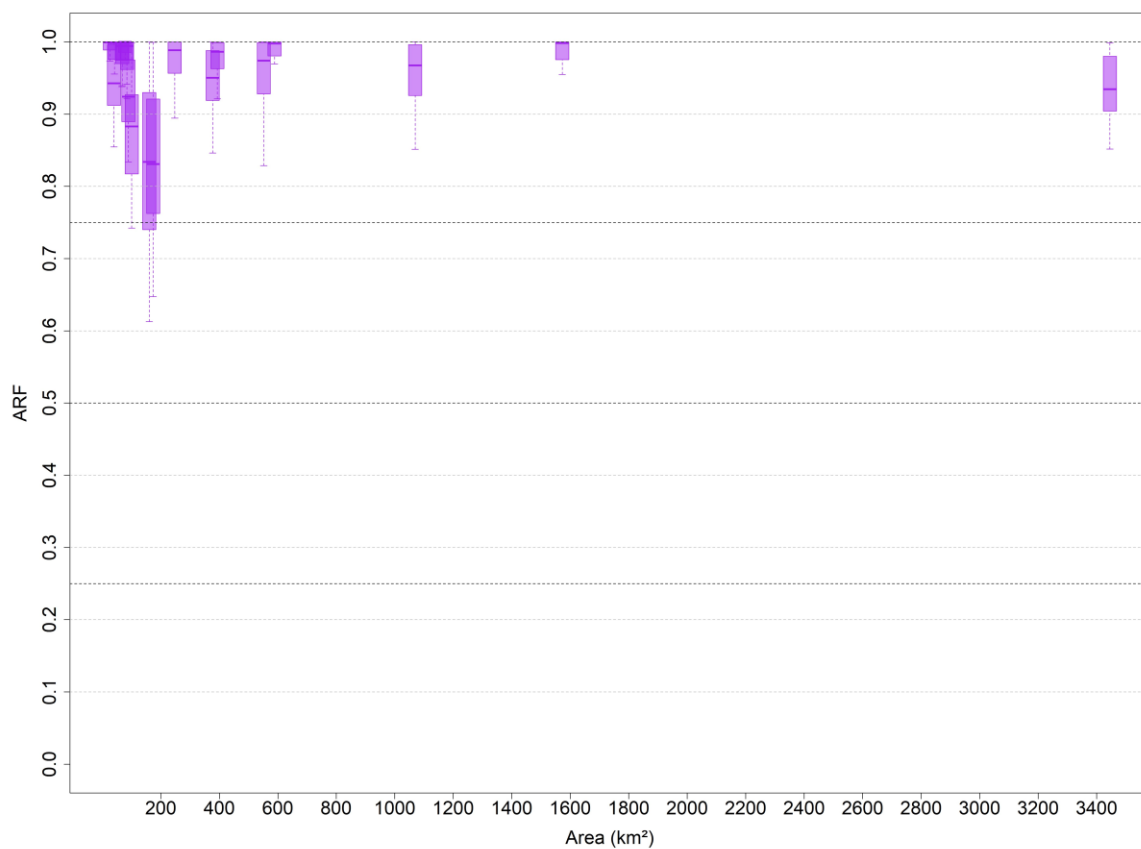


Figura 4.4 Boxplot dell'ARF medio nei 32 anni considerati, per il bacino del Reno e relativi sottobacini, fissata la durata di 1 (a), 3 (b), 6 (c), 12 (d), 24 (e), 36 (f), 48 (g).

estensione, mentre l'ampiezza della distribuzione incrementa per aree maggiori. Inoltre, data la tendenza generale dell'ARF a crescere con l'aumentare della durata, attestandosi intorno al valore unitario, si osserva un progressivo restringimento della distribuzione stessa verso un ARF pari a 1.

Se, piuttosto che far riferimento all'intera distribuzione, ci si concentra a considerare l'intervallo tra il 25° e il 75° percentile, detto intervallo interquartile, IQR, è possibile fornire una misura di quanto siano aggregati o distribuiti i dati intorno alla mediana, senza l'influenza dei valori estremi.

In generale, comunque, preme sottolineare la spiccata variabilità dei valori sperimentali annuali ottenuti per il fattore di ragguglio. La rappresentatività complessiva dell'ARF medio annuale risulta quindi piuttosto limitata, indagini future svolgeranno uno studio mirato della dipendenza della variabilità dei valori sperimentali del fattore dalle caratteristiche morfologiche dei bacini considerati, da quelle climatiche, oltre che dalla frequenza probabile dell'evento meteorico che ha prodotto il valore massimo annuale di precipitazione areale.

Area (km ²)	25,6	39,5	41	66,5	67,7	77,5	83,5	84	88,7	100,2	160,4	174	247,3	377,2	393,1	551,8	588,7	1069,5	1572,6	3445,3
d = 1h	0,1	0,1	0,09	0,1	0,1	0,11	0,22	0,18	0,13	0,132	0,161	0,11	0,19	0,164	0,163	0,172	0,164	0,119	0,13	0,109
d = 3h	0,05	0,13	0,05	0,05	0,05	0,08	0,1	0,09	0,11	0,129	0,129	0,16	0,084	0,112	0,063	0,129	0,1	0,121	0,049	0,099
d = 6h	0,02	0,14	0,04	0,03	0,04	0,04	0,06	0,06	0,12	0,133	0,178	0,11	0,115	0,149	0,07	0,094	0,082	0,099	0,044	0,11
d = 12h	0,02	0,13	0,04	0,03	0,04	0,03	0,04	0,05	0,09	0,1	0,233	0,14	0,081	0,114	0,039	0,145	0,065	0,096	0,038	0,088
d = 24h	0,03	0,08	0,03	0,01	0,03	0,03	0,03	0,04	0,09	0,115	0,164	0,16	0,039	0,071	0,035	0,09	0,041	0,1	0,032	0,081
d = 36h	0,01	0,13	0,02	0,01	0,03	0,03	0,04	0,04	0,08	0,125	0,135	0,13	0,03	0,124	0,033	0,069	0,025	0,091	0,022	0,072
d = 48h	0,01	0,08	0,02	0,03	0,03	0,02	0,02	0,04	0,09	0,11	0,19	0,16	0,043	0,069	0,036	0,071	0,019	0,07	0,025	0,076

Tabella 4.2 Tabelle riassuntive delle statistiche rappresentate dai boxplot, per ciascuna area considerata e ciascuna durata.

I risultati in Tabella rafforzano le osservazioni finora svolte: è evidente che per aree ridotte i valori dell'ARF annuale tendono a restringersi attorno alla mediana, mentre per aree di maggiore estensione si osserva una dispersione più alta; per quanto riguarda la durata, si osserva l'effetto contrario: durate ridotte dell'evento di pioggia comportano una maggiore variabilità, mentre per durate più elevate la distribuzione è ristretta attorno ad un valore prossimo all'unità.

4.2 Calibrazione di alcune leggi interpretative del fattore di riduzione areale ARF(A, d) sui bacini esaminati

Fissata la durata d, si è considerato di riportare sullo stesso piano (ARF, A) gli ARF medi annuali per la durata d ricavati per ciascun bacino.

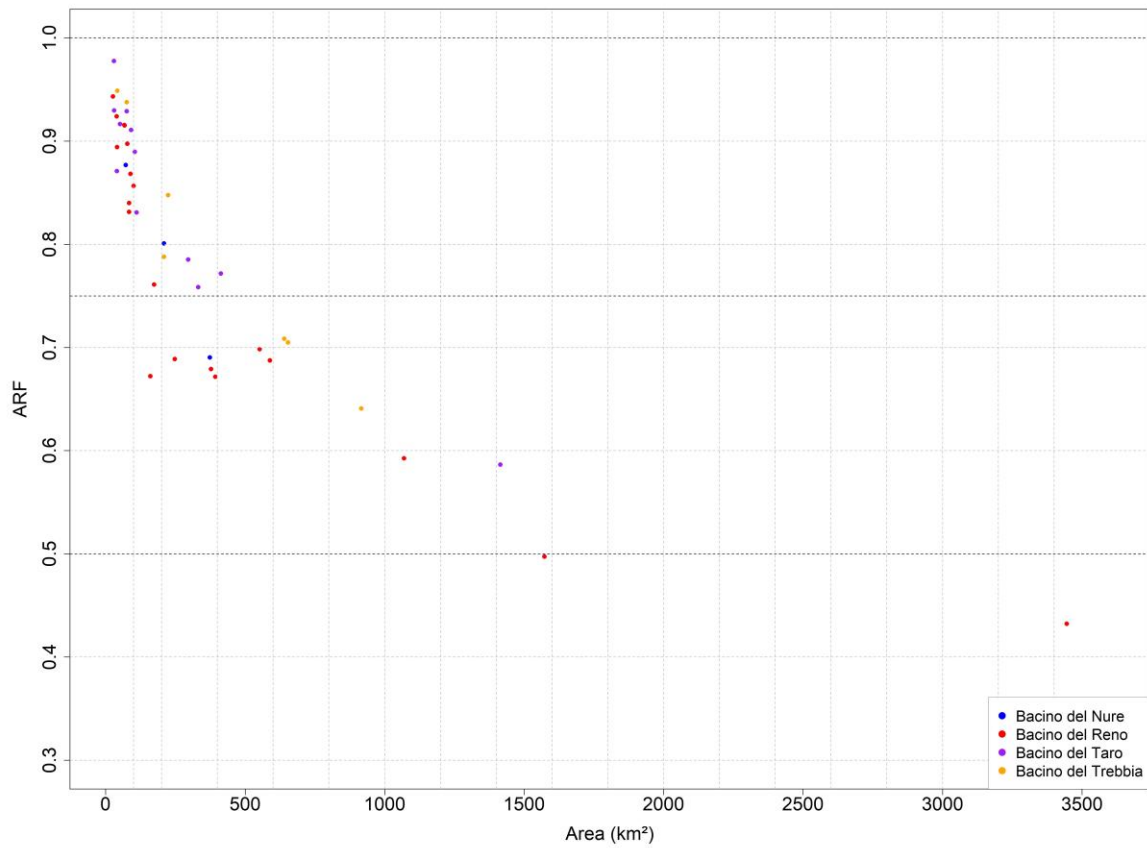
Dalle rappresentazioni ottenute per ciascuna durata, è possibile osservare un comportamento comune, analogo, in linea generale, a quanto descritto per il bacino del Reno e relativi sottobacini nel paragrafo precedente. Si è quindi considerato ragionevole indagare se l'ARF ottenuto dall'elaborazione delle osservazioni può essere ben stimato da alcune formule di letteratura, nel dettaglio le espressioni proposte dagli studi dell'USWB, di Moisello e Papiri e di Columbo; su un campione di osservazioni costituito dagli ARF medi annuali per ciascuna durata e per ogni bacino preso in esame, riferiti al periodo di studio, 1991-2022. In particolare, si fa riferimento al valor medio annuale del fattore di riduzione della precipitazione all'area poiché in letteratura si ritiene valido ragionare in tali termini. Nel calcolo dei valori sperimentali dell'ARF medio annuale, si è trascurata la dipendenza dal tempo di ritorno T e quindi dall'entità dell'altezza di precipitazione, data l'irrelevante influenza nella stima dell'ARF e in accordo gli Autori presi a riferimento in sede di taratura. Precisamente, la formulazione considerata con riferimento agli studi di Columbo non corrisponde a quella ottenuta direttamente dall'autore, bensì un'espressione che è stata ricavata, sotto le indicazioni dell'autore stesso, per estendere le sue osservazioni ad aree superiori a quelle contemplate dal metodo da lui proposto.

La calibrazione delle diverse relazioni è stata realizzata tramite un algoritmo di ottimizzazione, che trova il set di parametri in corrispondenza del quale si ottiene il massimo valore della funzione obiettivo scelta, in questo caso l'efficienza di Nash-Sutcliffe modificata.

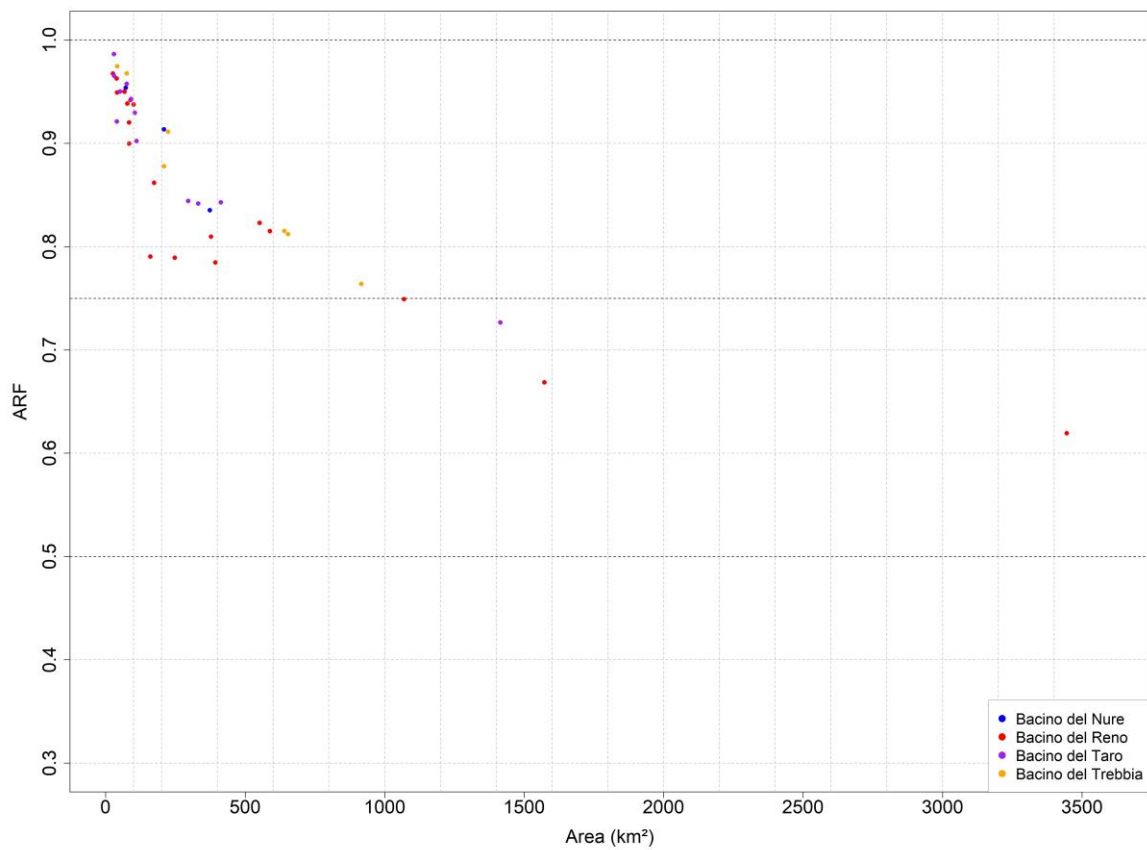
Dopo una breve descrizione delle metriche considerate, si riportano i risultati delle calibrazioni dei parametri caratteristici delle tre equazioni.

Di seguito è riportato, per ciascuna durata, lo scatterplot dell'intero campione di osservazioni, in Figura 4.5 a-g.

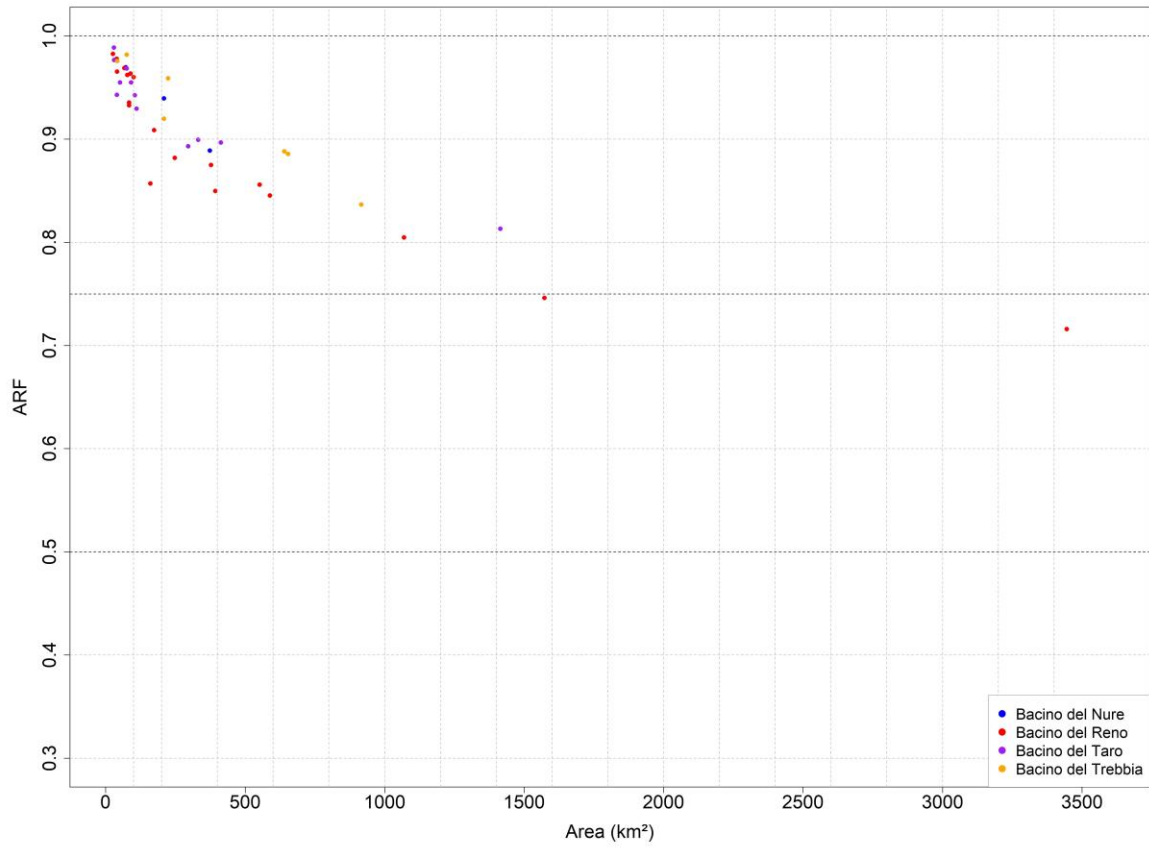
Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 1 h



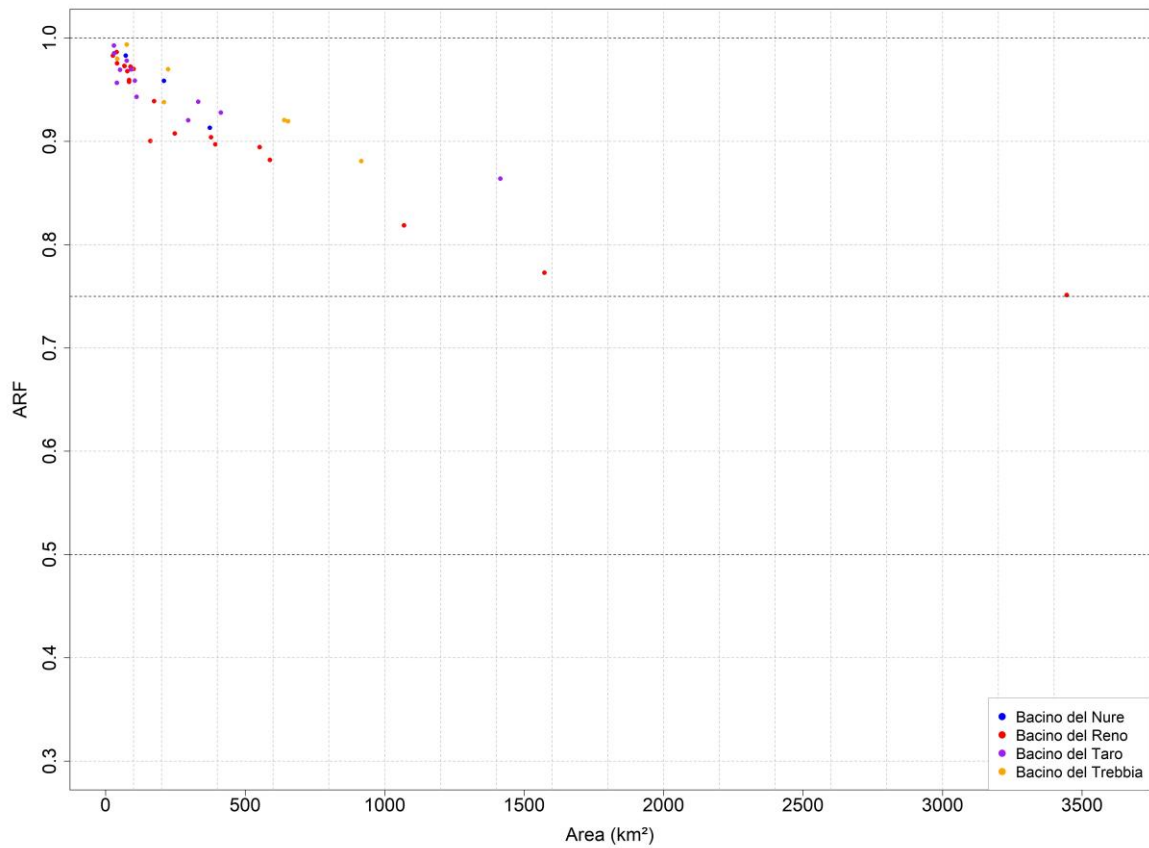
Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 3 h



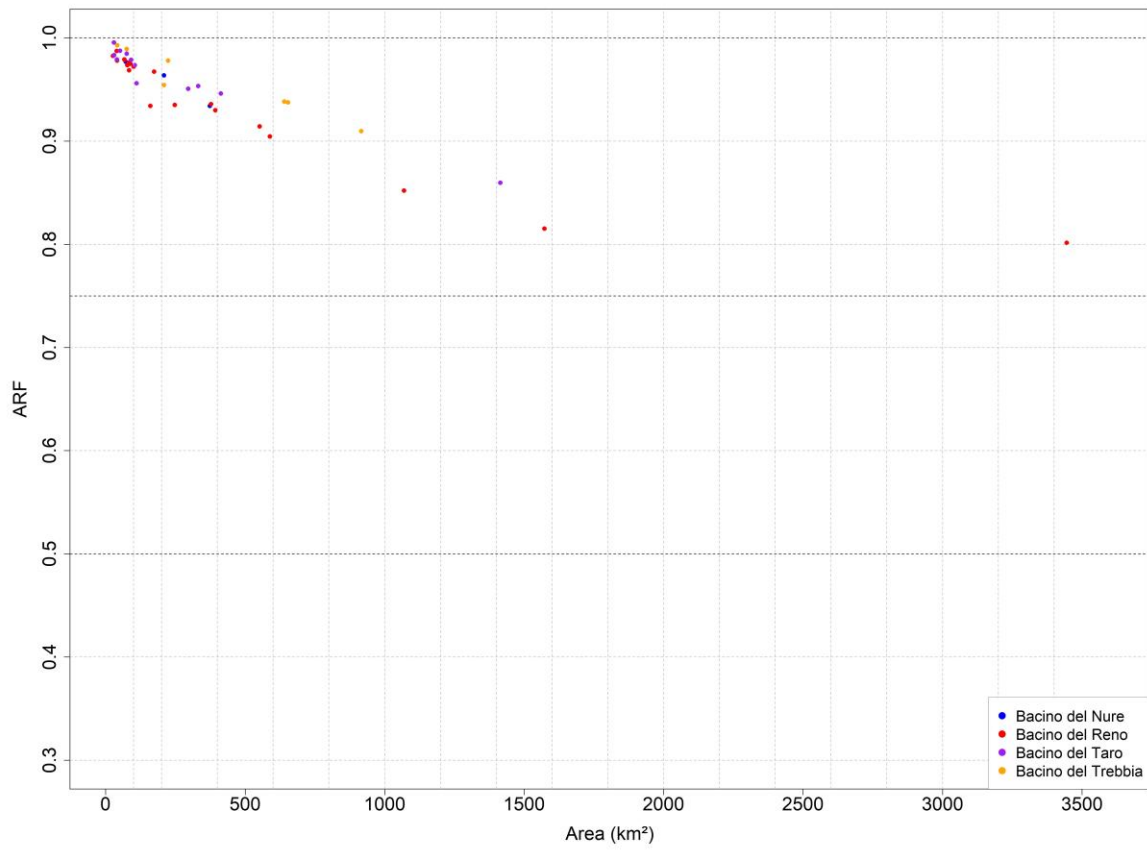
Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 6 h



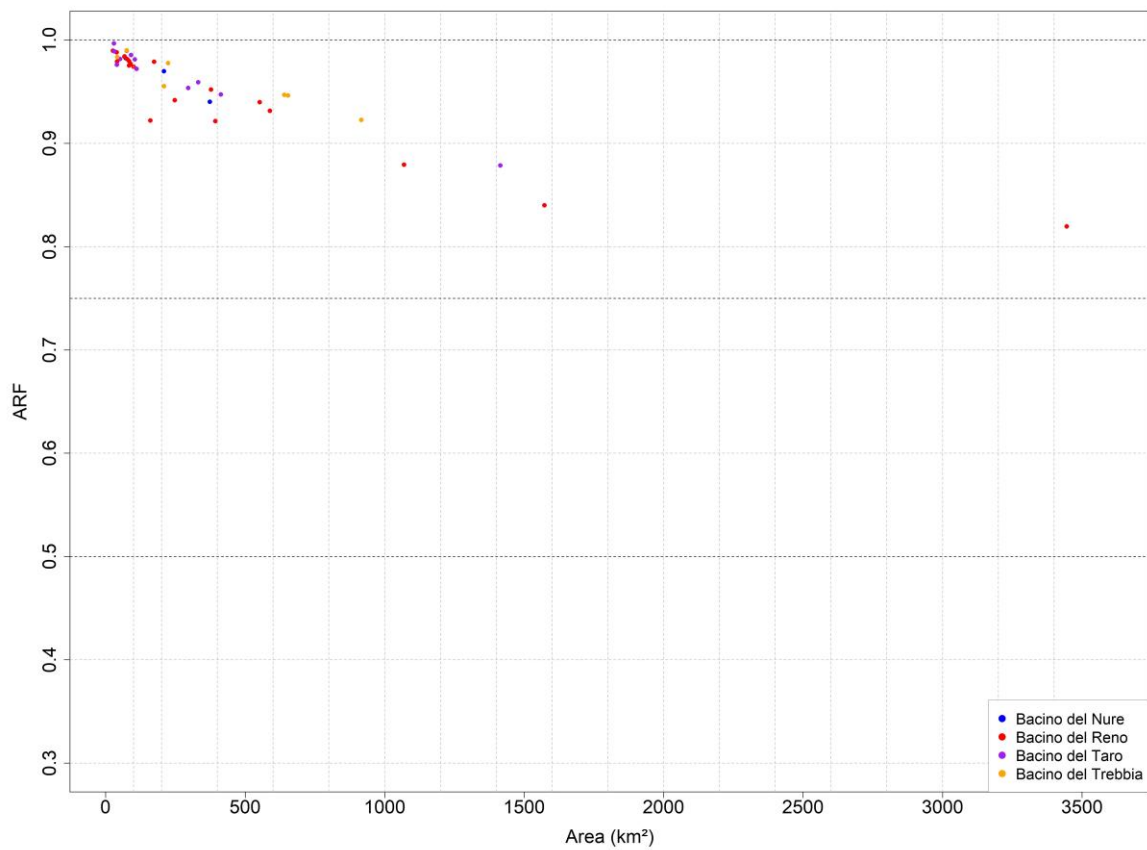
Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 12 h



Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 24 h



Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 36 h



Andamento dell'ARF medio negli anni in funzione dell'area, per i bacini appenninici presi in esame e una durata di 48 h

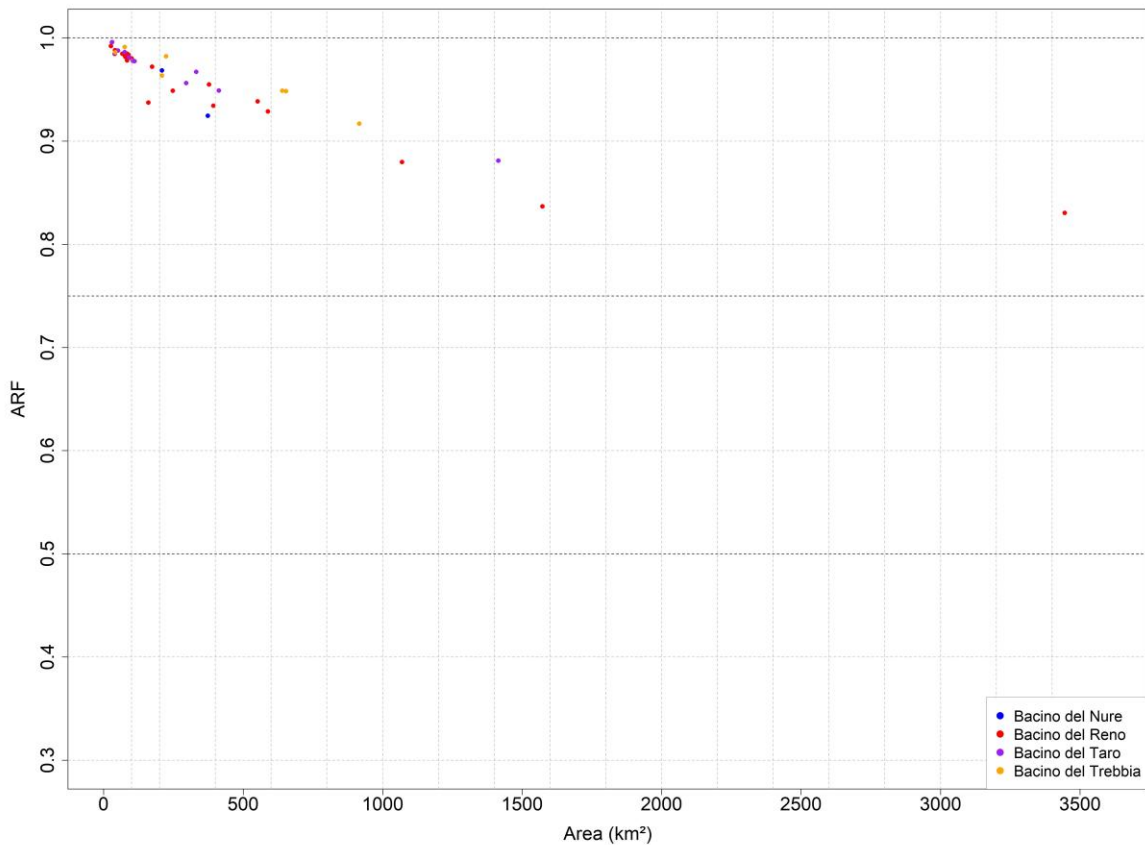


Figura 4.5 Scatterplot dell'ARF medio nei 32 anni considerati, in funzione dell'area, per il bacino appenninici considerati e relativi sottobacini, fissata la durata di 1 (a), 3 (b), 6 (c), 12 (d), 24 (e), 36 (f), 48 (g).

4.2.1 Metriche di prestazione considerate

Al fine di poter confrontare opportunamente i risultati ottenuti dalla calibrazione delle diverse relazioni, si è scelto di impiegare metriche che consentano di valutare l'efficienza dei modelli esaminati, tenendo conto del numero di parametri e di osservazioni a disposizione.

Le funzioni di costo impiegate sono, nel dettaglio, le seguenti misure di efficienza: il coefficiente R-quadro corretto e l'efficienza di Nash e Sutcliffe modificata. Si è inoltre valutata una misura di errore, il BIAS relativo.

4.2.1.1R² modificato

Il coefficiente R² riflette la bontà di adattamento di un modello alla variabile che intende spiegare, in termini di proporzione della varianza totale della variabile spiegata dalla regressione:

$$R^2 = \left(\frac{\text{cov}(x, \hat{x})}{\sigma_x \sigma_{\hat{x}}} \right)^2$$

Si tratta del quadrato del coefficiente di correlazione lineare tra osservazione e simulazione, o coefficiente di Pearson; più il suo valore è vicino a 1, maggiore è l'adattamento del modello alla variabile; viceversa, se la metrica tende a 0, il modello sarà meno affidabile.

Il coefficiente R^2 , d'altra parte, non penalizza l'inclusione di variabili esplicative non significative, al contrario tende ad assumere valori più elevati all'incrementare del numero di variabili indipendenti utilizzate per predire l'output del modello stesso, anche se potrebbe non esserci anche una reale crescita del valore predittivo del modello.

Per questo motivo si ricorre al coefficiente di determinazione corretto:

$$R_{adj}^2 = 1 - \frac{N-1}{N-p} (1 - R^2)$$

Dove:

- N è la dimensione del campione;
- p è il numero dei parametri.

È evidente che più alti sono i valori di p , più lontano sarà l' R^2 corretto dal coefficiente standard; viceversa, a valori inferiori di p , la frazione centrale sarà vicina a 1 e, quindi, l' R^2 modificato e l' R^2 normale saranno più simili.

Bisogna però considerare che l' R^2 , in entrambe le espressioni, standard e corretta, vede solo la linearità tra ciò che simula e ciò che osserva, ovvero calcola la dispersione delle coppie (ARF_{oss} , ARF_{sim}) rispetto alla retta d'inclinazione $R^2=1$, non tenendo conto degli scarti: non vede distorsioni, bias ed errori sistematici. Di conseguenza, per valutare il funzionamento del modello si deve prendere qualche altra statistica.

4.2.1.2 Adjusted NSE

L'efficienza di Nash-Sutcliffe, anche detta varianza spiegata, è, per l'appunto, una misura di efficienza, per cui più elevato è il suo valore meglio va il modello, ma si tratta semplicemente di una adimensionalizzazione della sommatoria degli scarti al quadrato:

$$NSE = 1 - \frac{SSQ}{(N-1)var(x)} = 1 - \frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

Dove:

- SSQ indica la sommatoria degli scarti al quadrato;
- N è il numero di valori osservati nella serie temporale che si vuole riprodurre;
- $\text{var}(x)$ è la varianza campionaria;
- \bar{x} è il valor medio dei dati sperimentali.

Essendo una sommatoria degli scarti al quadrato, normalizzata rispetto al numero e alla varianza delle osservazioni, vede gli errori sistematici e le distorsioni, al contrario di R^2 , e fornisce una valutazione delle prestazioni di un modello che non è influenzata dall'entità del campione di osservazioni su cui viene calibrato il modello.

Così come per R^2 , è possibile fare riferimento all'efficienza di Nash-Sutcliffe corretta, NSE_{adj} , per la selezione del modello più adatto. Questa, infatti, tiene conto del numero di parametri del modello p , in modo tale da poter confrontare le prestazioni della stima delle osservazioni di modelli più complessi, cioè con un numero più elevato di parametri, con la performance di modelli più semplici:

$$NSE_{adj} = 1 - \frac{N - 1}{N - p} (1 - NSE)$$

Nell'espressione standard del coefficiente di Nash, l'unico termine sul quale giocano un ruolo i parametri del modello è il valore simulato della grandezza osservata \hat{x}_i presente al numeratore.

Il coefficiente moltiplicativo:

$$k = \frac{N - 1}{N - p}$$

Che appare nell'espressione dell' R^2 modificato e di NSE_{adj} , ha al numeratore il numero di osservazioni diminuito di 1, al fine di correggere la dispersione della varianza, quando la sua stima è basata su un numero ridotto di osservazioni. A denominatore vi sono i gradi di libertà del problema di calibrazione, ossia le osservazioni meno il numero dei parametri calibrati statisticamente.

4.2.1.3 Bias relativo

Infine, le prestazioni dei modelli di letteratura considerati sono state valutate anche considerando il bias relativo, una misura che indica quanto un valore stimato è deviato dal vero valore o dalla verità del parametro che si sta cercando di stimare. In altre parole, il bias relativo quantifica l'errore sistematico o il "bias" in una stima rispetto al valore vero o atteso, fornendo una misura del valor medio della distorsione relativa. La formula impiegata per calcolare il bias relativo è la seguente:

$$BIAS_{rel} = \frac{1}{J} \frac{1}{I} \sum_{j=1}^J \sum_{i=1}^I \frac{\hat{ARF}(A_j, d_i) - ARF(A_j, d_i)}{ARF(A_j, d_i)}$$

Un bias relativo pari a zero indica che la stima è perfettamente accurata e non presenta alcun errore sistematico. Si tratta di una misura della direzione e della magnitudine medie dell'errore sistematico nelle stime: un bias relativo positivo indica che la stima è sopravvalutata rispetto al vero valore, mentre un bias relativo negativo indica che la stima è sottostimata rispetto al vero valore.

4.1.1 Formula dell'USWB

La prima espressione tarata sui risultati ottenuti è quella proposta dall'U.S. Weather Bureau:

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

Dove A è l'area in km² e d è la durata dell'evento di pioggia in ore. I parametri ricalibrati sui bacini considerati sono i seguenti:

Parametri	Calibraz. NSE _{adj}
c ₁	-0,002431
c ₂	-0,8287
c ₃	0,2926

I valori trovati da Eagleson sono:

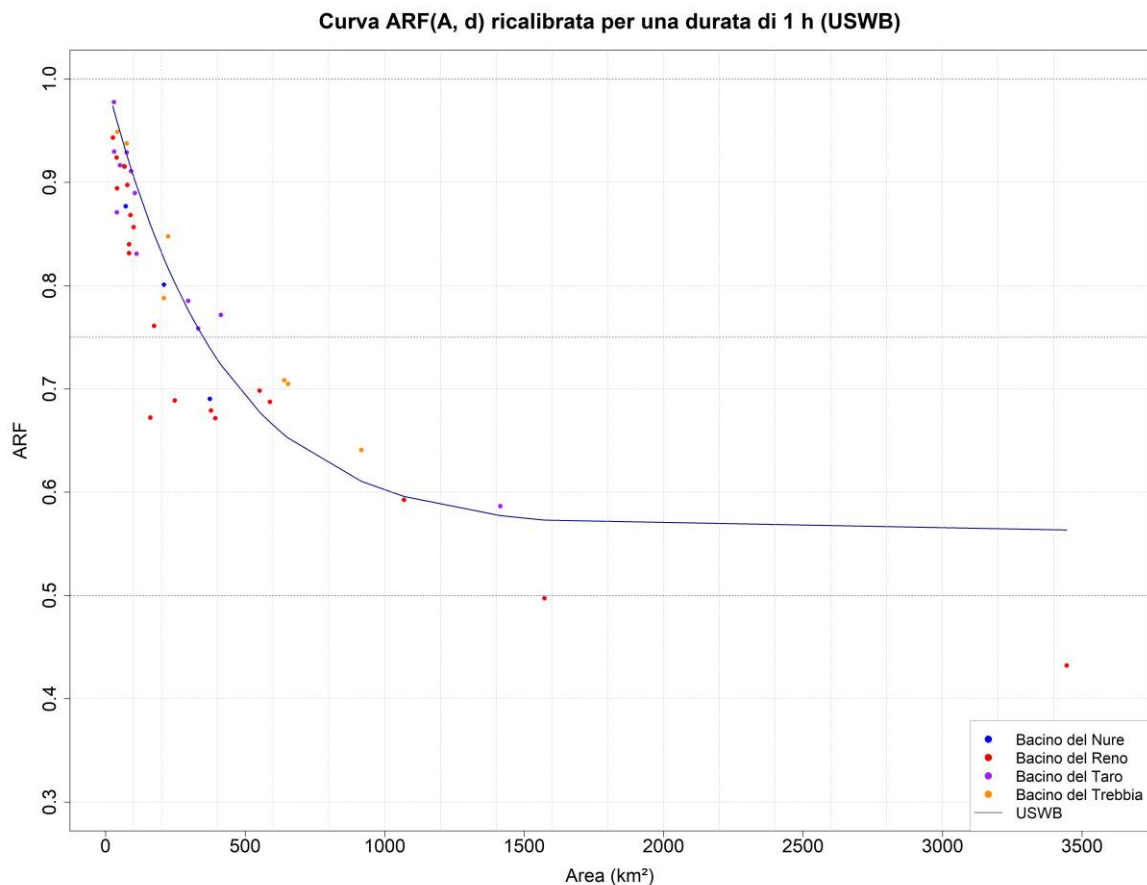
Parametri	Eagleson
c ₁	-0,00386
c ₂	-1,1
c ₃	0,25

In termini di metriche di prestazione, i valori ottenuti sono i seguenti:

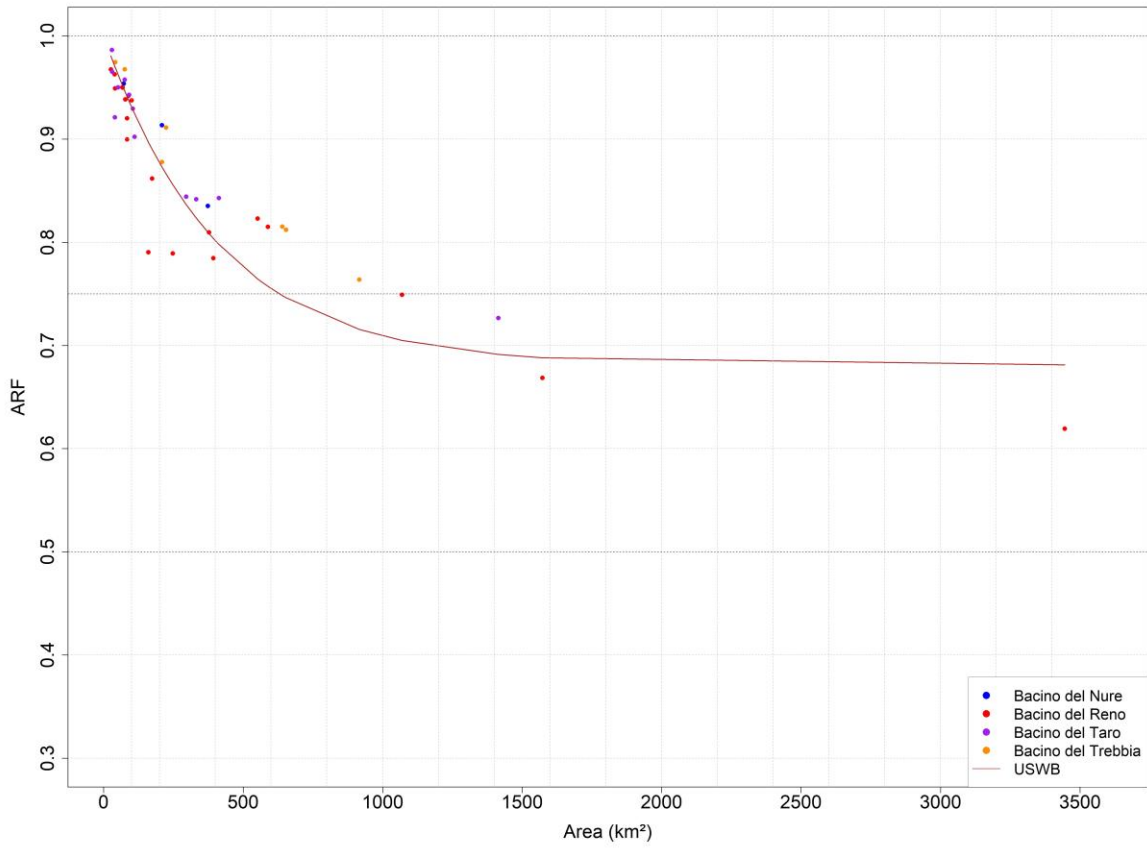
USWB	
SSQ	0,3405
NSE	0,8601
NSE _{adj}	0,8591
R ²	0,8630
R ² _{adj}	0,8620
BIAS _{rel}	0,00602

La prestazione del modello non è da ritenersi scarsa, considerando che, generalmente, si ritiene soddisfacente un valore dell'efficienza di Nash di 0,6-0,7, ovvero si ritiene adeguata la performance del modello quando questo spiega il 60-70% della variabilità delle osservazioni. Tuttavia, le restanti calibrazioni risultano migliori.

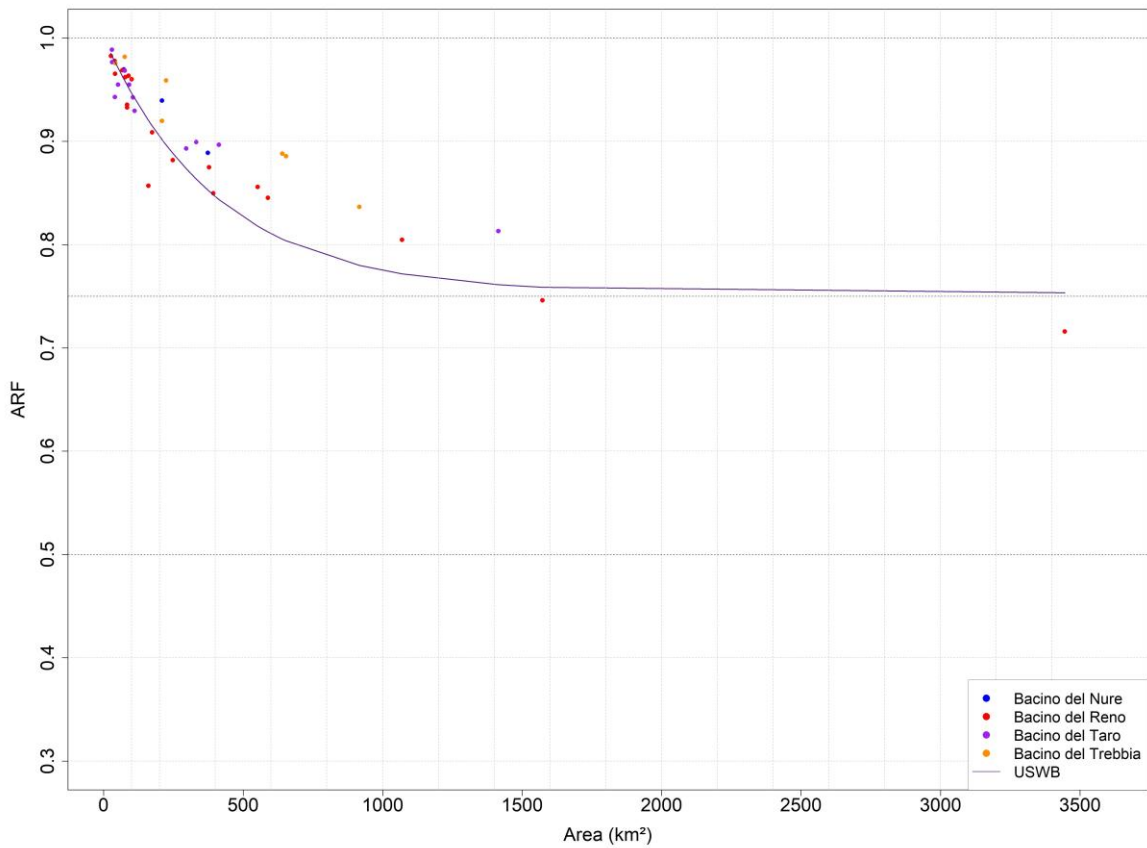
Si riportano i grafici contenenti i dati sperimentali e la calibrazione, per ciascuna durata, in Figura 4.6 a-g.



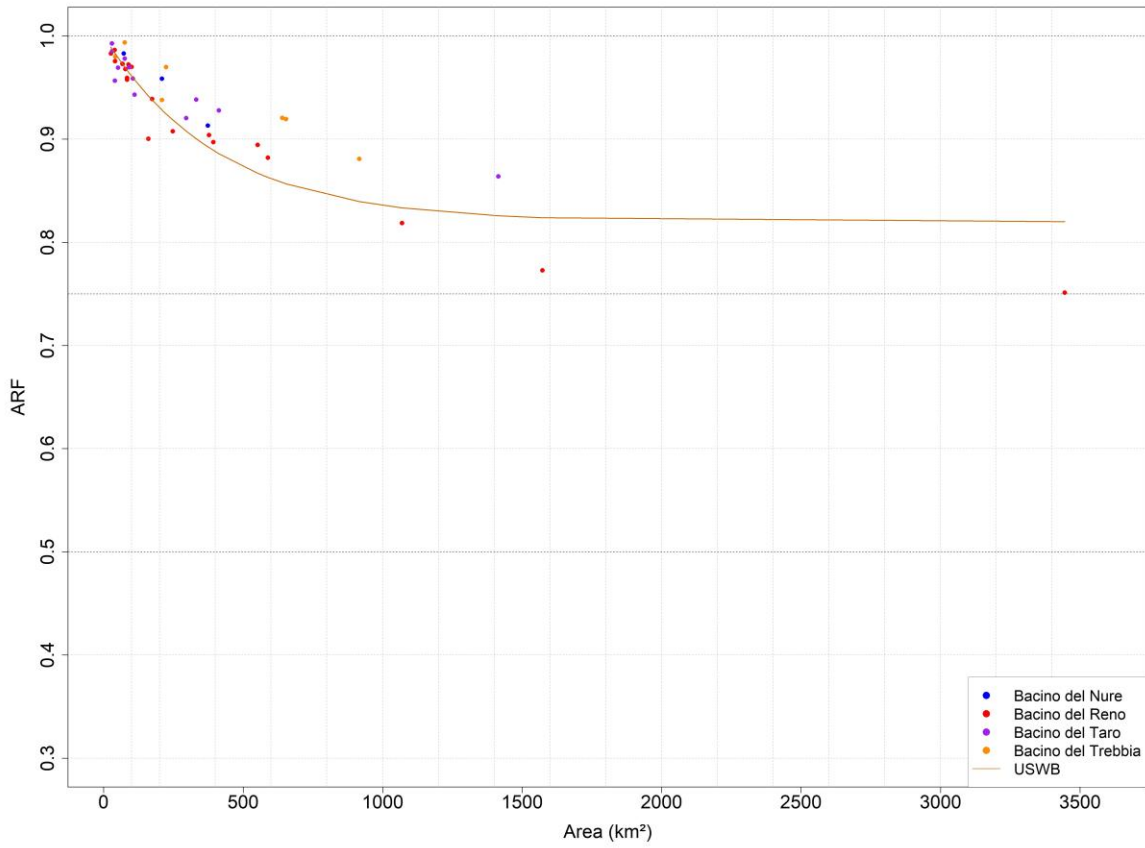
Curva ARF(A, d) ricalibrata per una durata di 3 h (USWB)



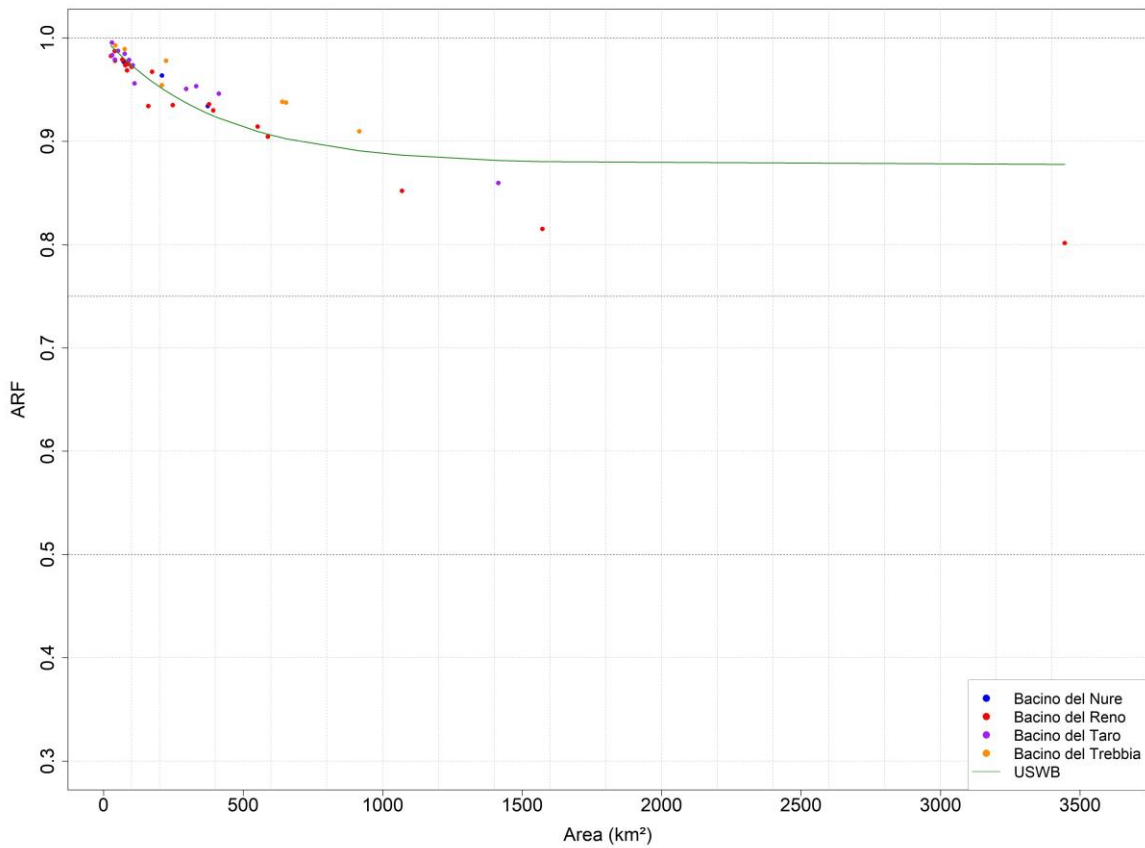
Curva ARF(A, d) ricalibrata per una durata di 6 h (USWB)



Curva ARF(A, d) ricalibrata per una durata di 12 h (USWB)



Curva ARF(A, d) ricalibrata per una durata di 24 h (USWB)



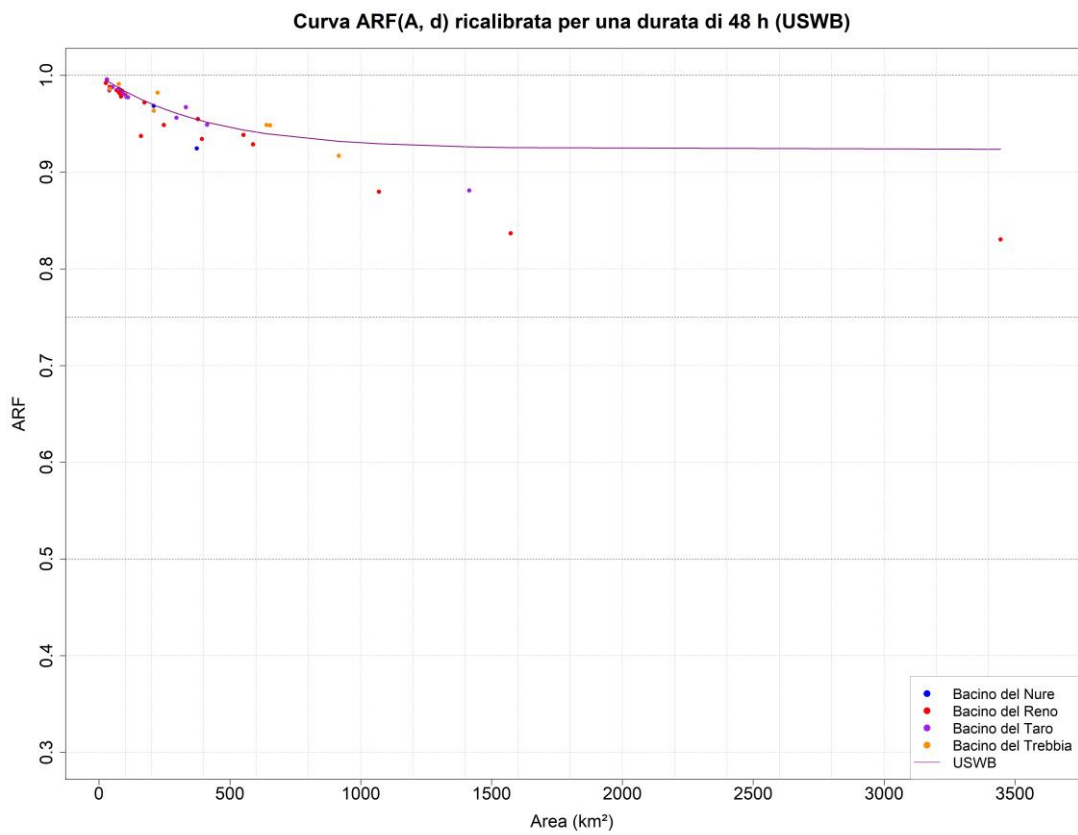
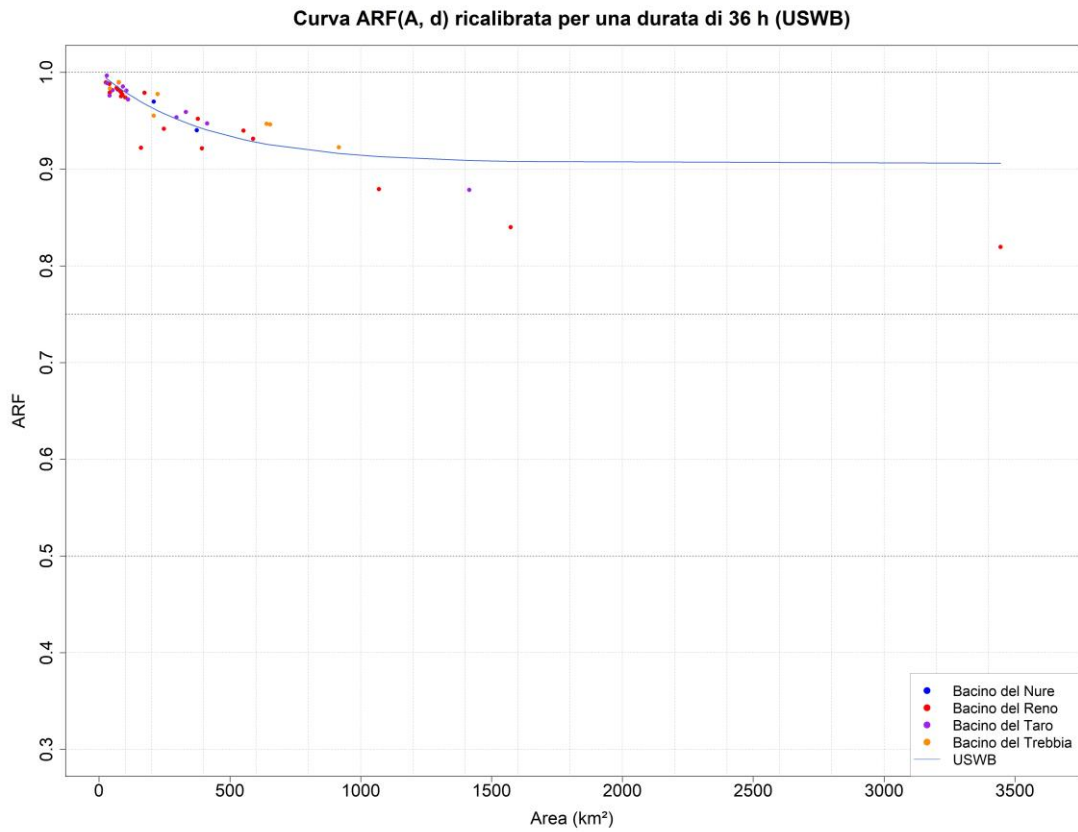


Figura 4.6 Curve ARF(A, d) descritte dalla formulazione dell'U.S. Weather Bureau ricalibrate su valori sperimentali ricavati sui bacini appenninici esaminati, fissata la durata di 1 (a), 3 (b), 6 (c), 12 (d), 24 (e), 36 (f), 48 (g).

La formulazione dell'USWB è stata ricavata per durate che vanno da 30 minuti a 24 ore e aree dai 250 fino a circa 1000 km²; tuttavia, si osserva un decadimento della prestazione della curva in corrispondenza dei valori massimi di area, circa 3500 km². Questo, si è immaginato, fosse dovuto non solo ai limiti del dominio della relazione proposta da Eagleson, ma anche alla scarsa presenza nel campione di ARF_Y di valori riferiti ad aree di grande estensione, a differenza di quanto si osserva per i valori intermedi di area, per i quali si ha a disposizione un elevato numero valutazioni sperimentali, che quindi hanno un maggiore peso in fase di calibrazione. La legge, quindi, si potrebbe ritenere adatta all'interpretazione dell'ARF sui bacini considerati nel campo di area già individuato dall'USWB. Per ciascuna durata, le prestazioni del modello ricavato in termini di NSE_{adj} sono le seguenti:

Durata (h)	NSE _{adj}
1	0,7819
3	0,8178
6	0,7314
12	0,7440
24	0,7709
36	0,6872
48	0,5943

In sede di calibrazione, infatti, le curve sono state tarate considerando l'intero set campionario di dati sperimentali, costituito dall'ARF annuale per ciascuna durata e ciascuna area considerata. Di conseguenza, le metriche presentate sono valutazioni globali della performance del modello ottenuto; si osserva che le prestazioni tendono a decadere con l'aumentare della durata, coerentemente ai limiti di durata dell'evento di pioggia considerati dall'Autore della stessa.

Si osserva quindi, elemento evidente anche dai grafici, che la performance varia sulla base della durata e delle aree prese in esame: per brevi durate c'è un forte scostamento con l'ARF sperimentale relativo all'area massima considerata, di circa 3500 km².

Confrontando la formulazione di letteratura con quella ricalibrata sui valori raccolti, rappresentandone gli andamenti sul grafico ARF-A nella Figura 4.7:

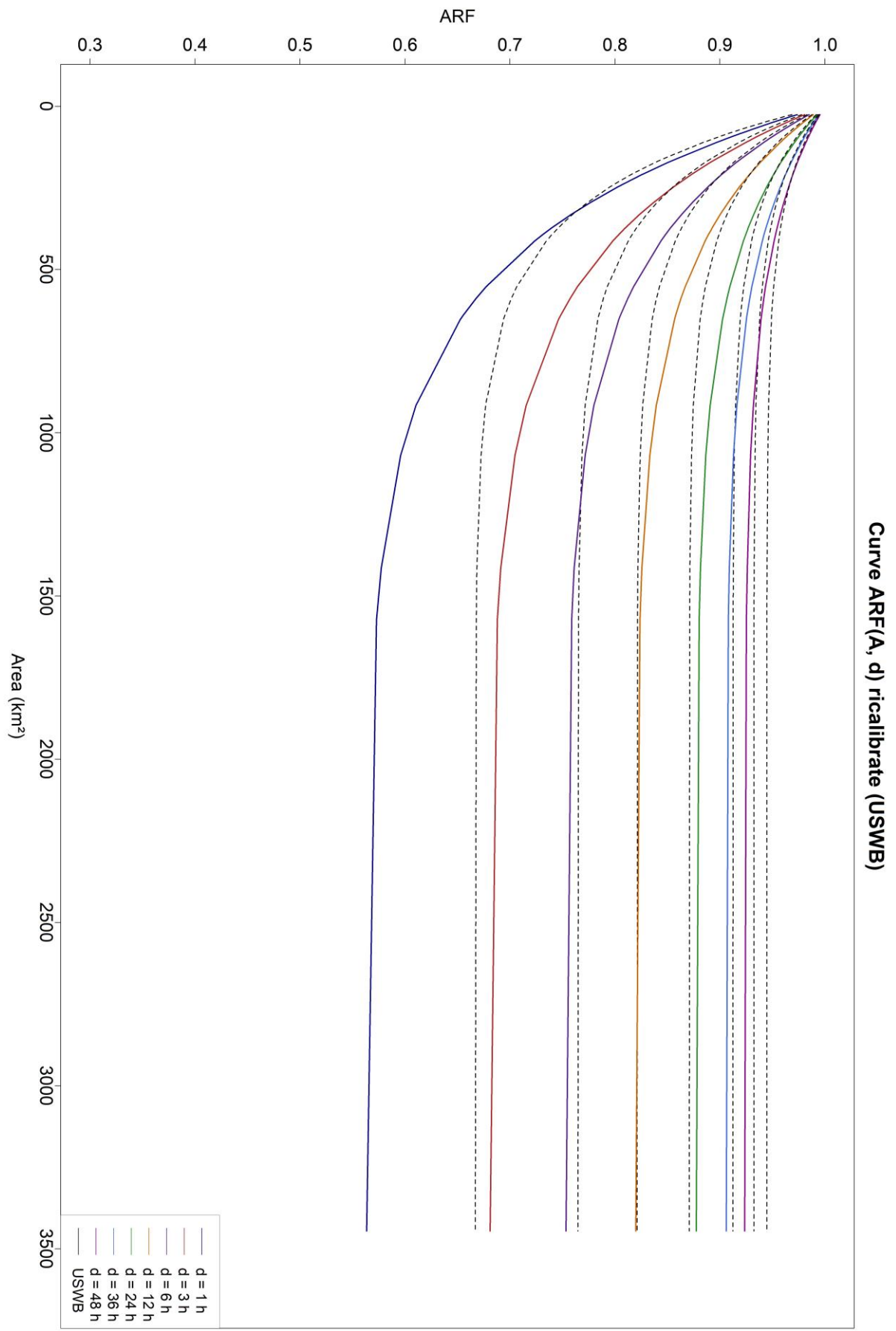


Figura 4.7 Confronto con la curva ottenuta da Eagleson.

Le curve di letteratura si discostano molto dalle curve ricalibrate sui bacini appenninici considerati, perché formulate su una zona climaticamente diversa; in particolare, queste tendono a fornire una forte sovrastima del fattore di riduzione della precipitazione all'area.

In conclusione, l'espressione della relazione proposta a seguito della calibrazione della relazione dell'USWB è la seguente:

$$ARF(A, d) = 1 - (1 - \exp(-0,0024A))\exp(-0,83d^{0,29})$$

4.1.2 Formula di Moisello e Papiri

La formula proposta da Moisello e Papiri, in generale:

$$ARF(A, d) = 1 - \exp[c_1 A^{c_2} d^{(c_3 - \exp(c_4 A^{c_5}))}]$$

Con A in km² e d in ore, è stata ricavata dagli Autori per precipitazioni di durata inferiore alle 12 ore e aree fino a 800 km². I coefficienti da loro ottenuti sono i seguenti:

Parametri	Moisello e Papiri
c ₁	-2,472
c ₂	-0,242
c ₃	0,6
c ₄	-0,643
c ₅	0,235

Mentre quelli ricavati per i bacini appenninici considerati:

Parametri	Calibraz. NSE _{adj}
c ₁	-7,8792
c ₂	-0,2988
c ₃	0,2603
c ₄	-1,0131
c ₅	0,2174

È interessante evidenziare la differenza tra il parametro c₁ calibrato da Moisello e Papiri e quello ottenuto dalla taratura. Riportando le curve di letteratura e quelle ottenute dalla presente analisi sullo stesso piano (Figura 4.8), si ha:

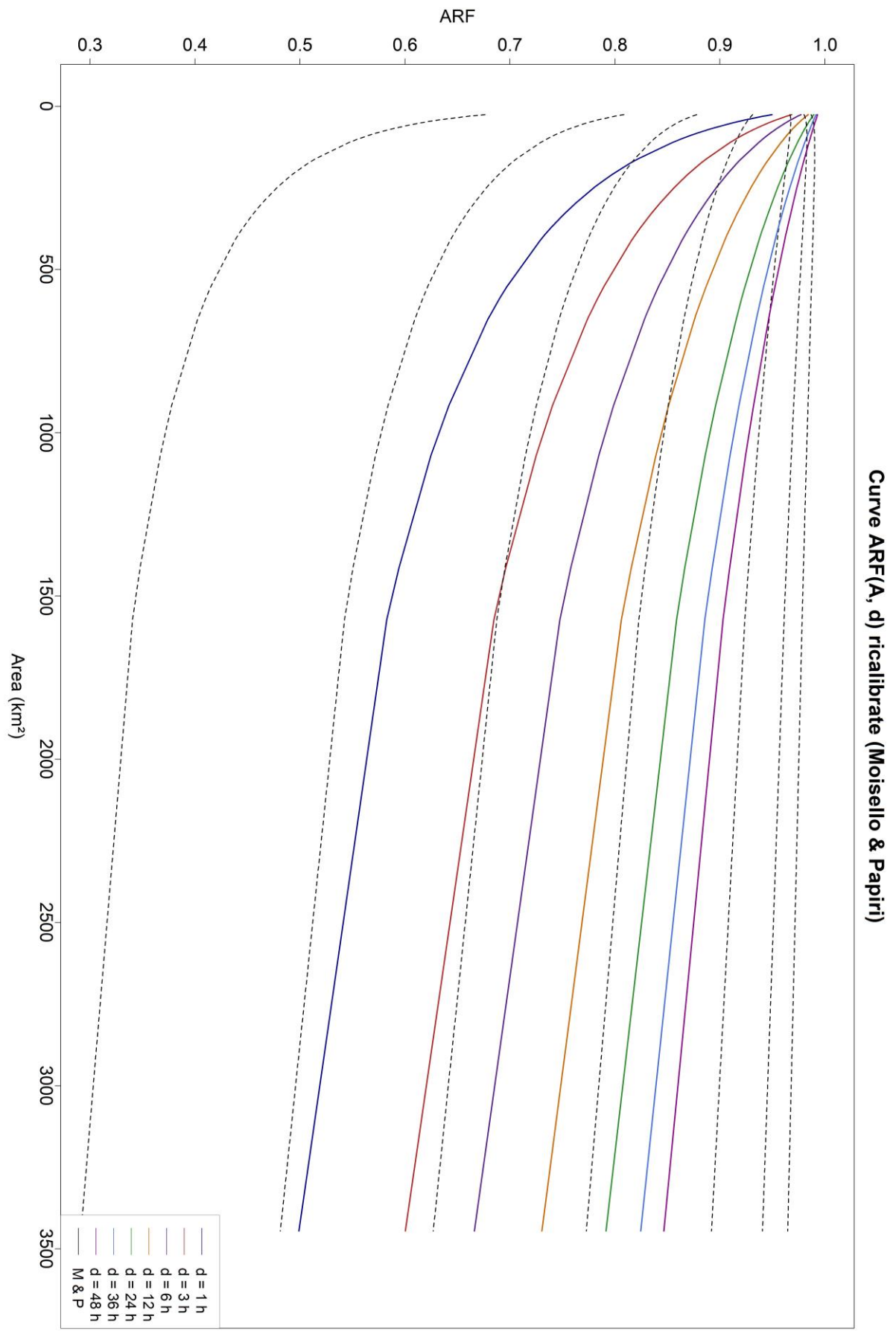
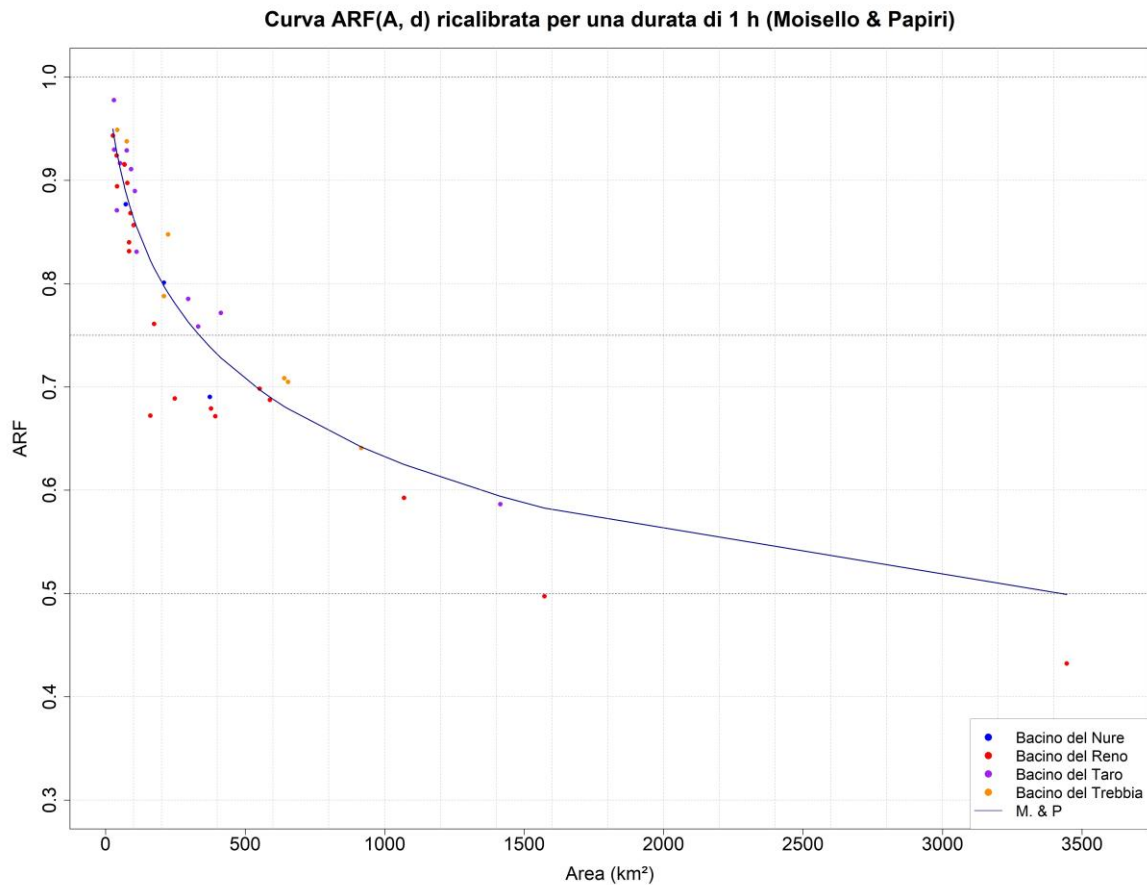
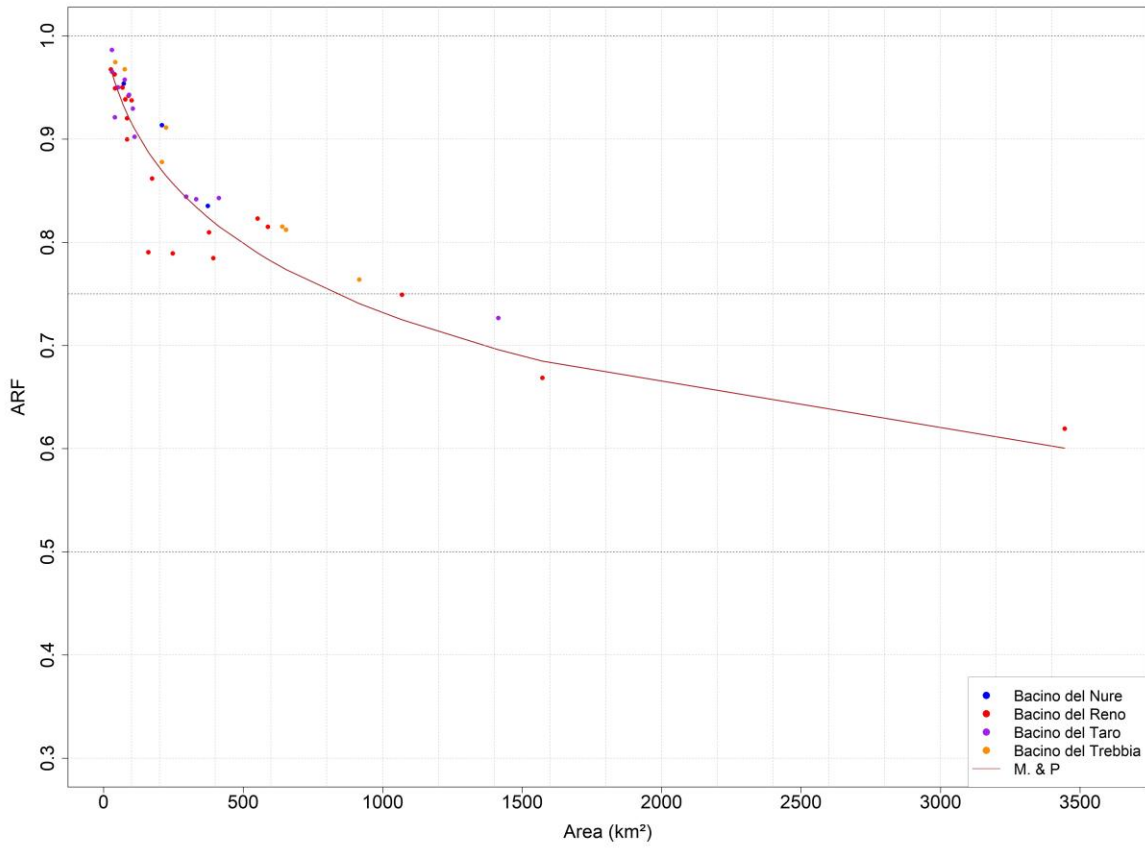


Figura 4.8 Confronto con la curva ottenuta da Moiseello e Papiri.

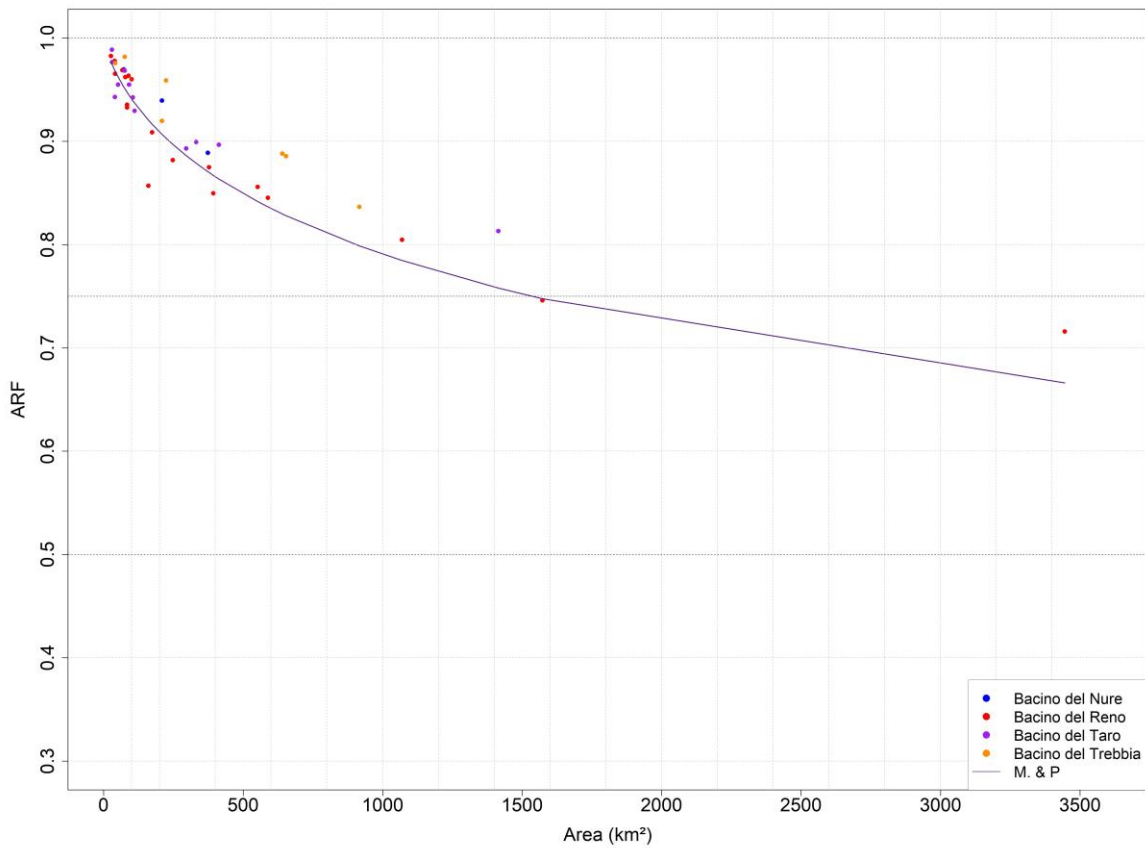
La curva proposta da Moisello e Papiri sottostima l'ARF per brevi durate (1 e 3 h) e lo sovrastima per le restanti, mentre è sufficientemente rappresentativa solo per le 6 ore. Dalla rappresentazione delle curve fittate sui parametri in output alla calibrazione (Figura 4.9 a-g), si osserva che in linea generale la formulazione di Moisello e Papiri tarata sui bacini sub-appenninici tosco-emiliani appare molto adatta all'interpretazione del fenomeno di riduzione su tali aree, se ricalibrata sui valori osservati.



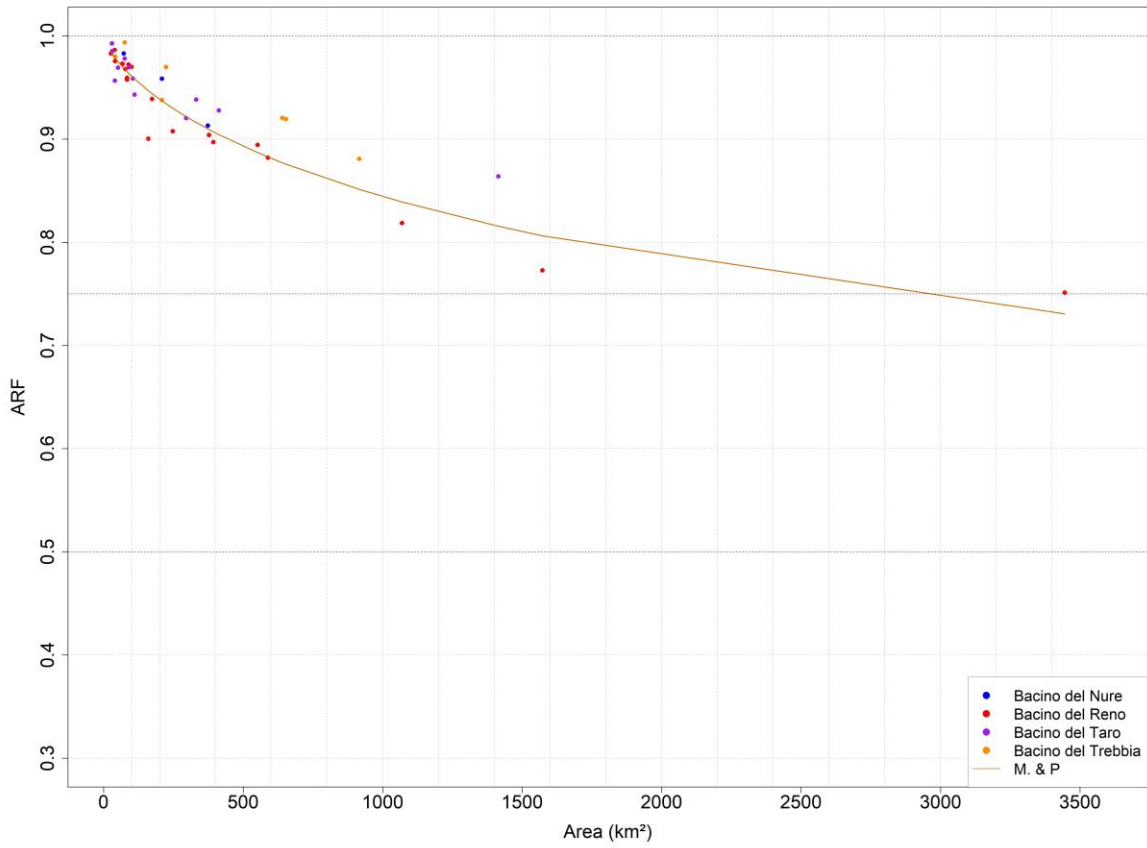
Curva ARF(A, d) ricalibrata per una durata di 3 h (Moisello & Papiri)



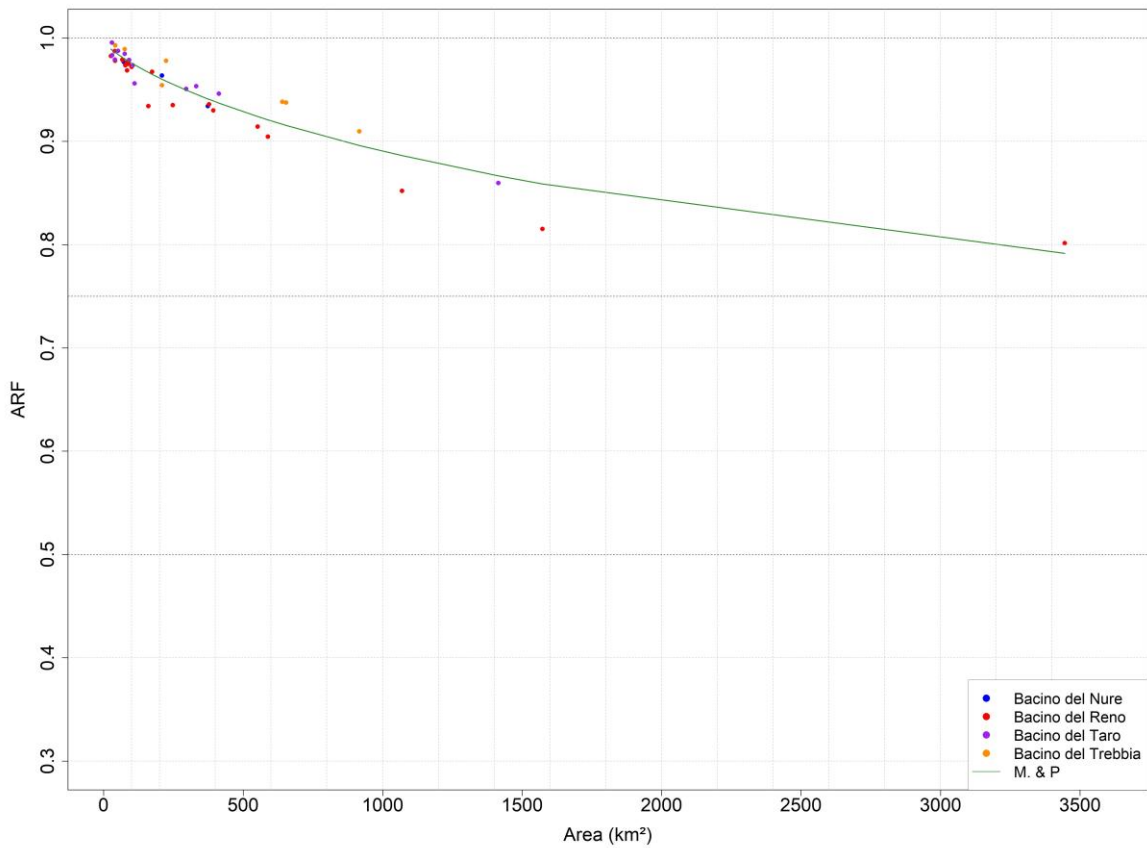
Curva ARF(A, d) ricalibrata per una durata di 6 h (Moisello & Papiri)



Curva ARF(A, d) ricalibrata per una durata di 12 h (Moisello & Papiri)



Curva ARF(A, d) ricalibrata per una durata di 24 h (Moisello & Papiri)



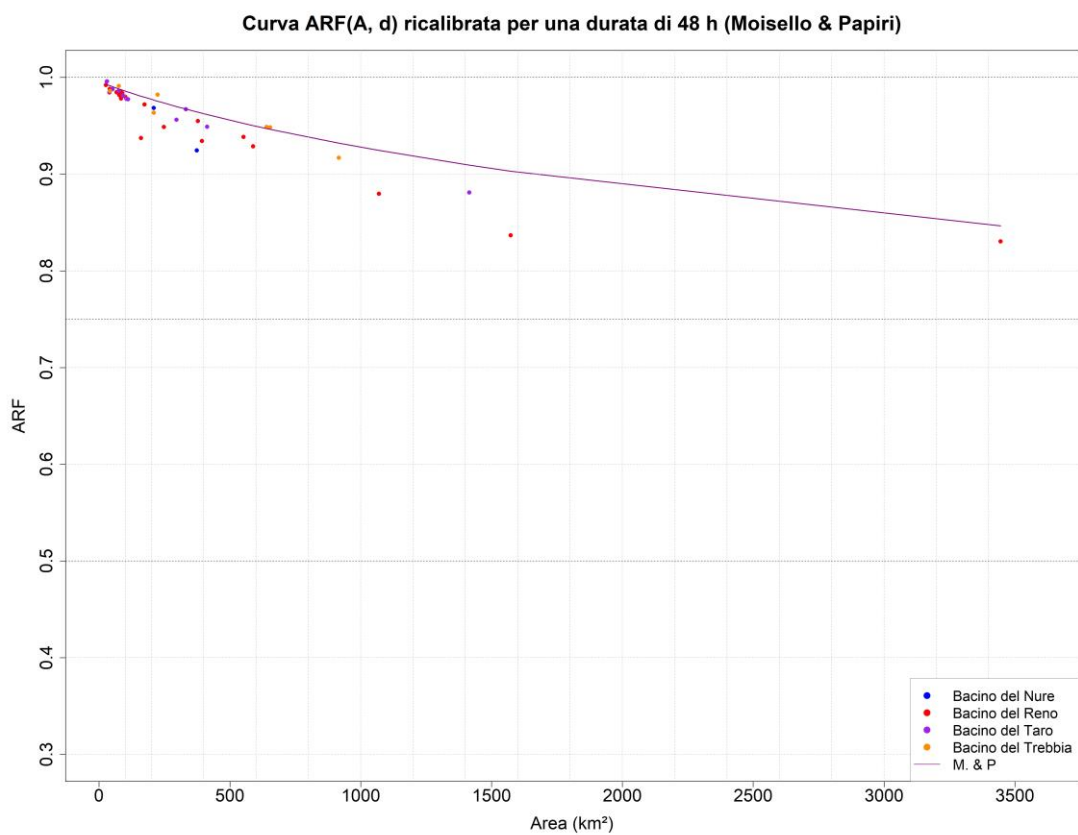
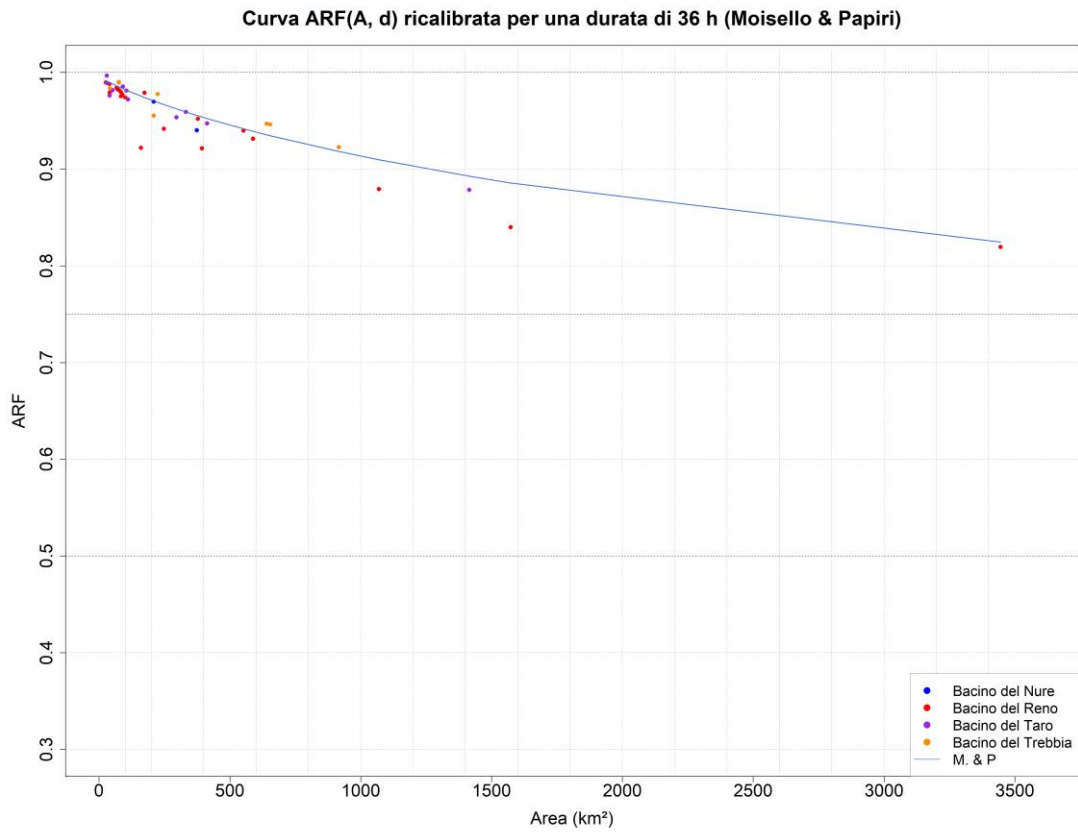


Figura 4.9 Curve ARF(A, d) descritte dalla formulazione proposta da Moisello e Papiri ricalibrate su valori sperimentali ricavati sui bacini appenninici esaminati, fissata la durata di 1 (a), 3 (b), 6 (c), 12 (d), 24 (e), 36 (f), 48 (g).

In termini di funzioni di costo, è stato ottenuto:

Moisello e Papiri	
SSQ	0,1973
NSE	0,9189
NSE _{adj}	0,9178
R ²	0,9191
R ² _{adj}	0,9180
BIAS _{rel}	0,00229

Mentre con riferimento a ciascuna durata, in termini di NSE_{adj}:

Durata (h)	NSE _{adj}
1	0,8701
3	0,8745
6	0,8107
12	0,8525
24	0,8988
36	0,8402
48	0,7547

Le prestazioni tendono a crescere all'aumentare della durata dell'evento di pioggia, fino alle 24 ore, oltre le quali decadono, risultato atteso, poiché le curve sono state stimate dagli Autori per durate fino alle 12 ore.

La formulazione di Moisello e Papiri ricalibrata sui bacini considerati:

$$ARF(A, d) = 1 - \exp \left[-7,88A^{-0,3} d^{(0,26 - \exp(-1,01A^{0,22}))} \right]$$

È quella che meglio rappresenta i valori sperimentali ottenuti, come è evidente dai valori delle metriche considerate, che rappresentano un risultato certamente soddisfacente.

4.1.3 Formula di Columbo

Infine, è stata ricalibrata, sulle osservazioni ricavate, la formula di Columbo, opportunamente modificata, che ha la seguente forma:

$$ARF(A, d) = (1 - c_1 A^{c_2}) d^{(c_3 A^{c_4})}$$

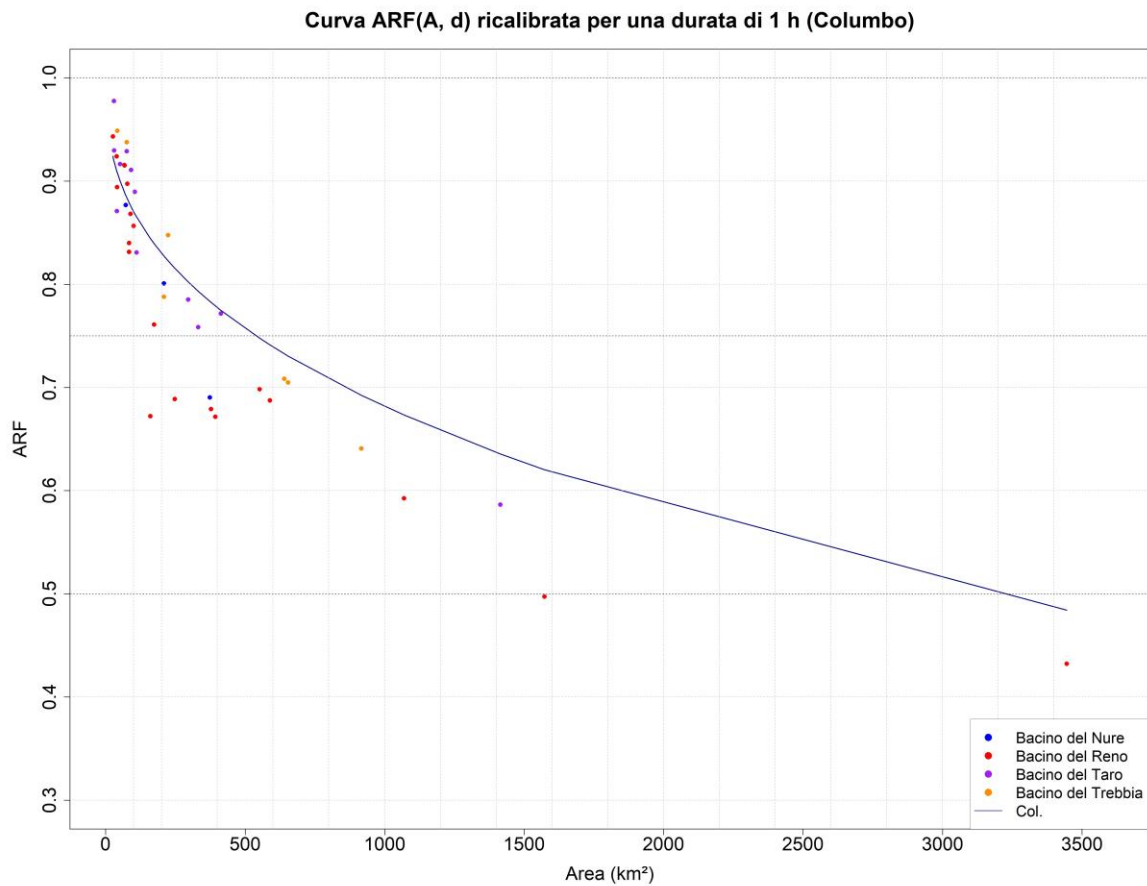
Dove A è l'area del bacino in km² e d è espressa in ore. I coefficienti ricalibrati sono:

Parametri	Calibraz. NSE _{adj}
c ₁	0,02142
c ₂	0,3906
c ₃	0,005447
c ₄	0,3946

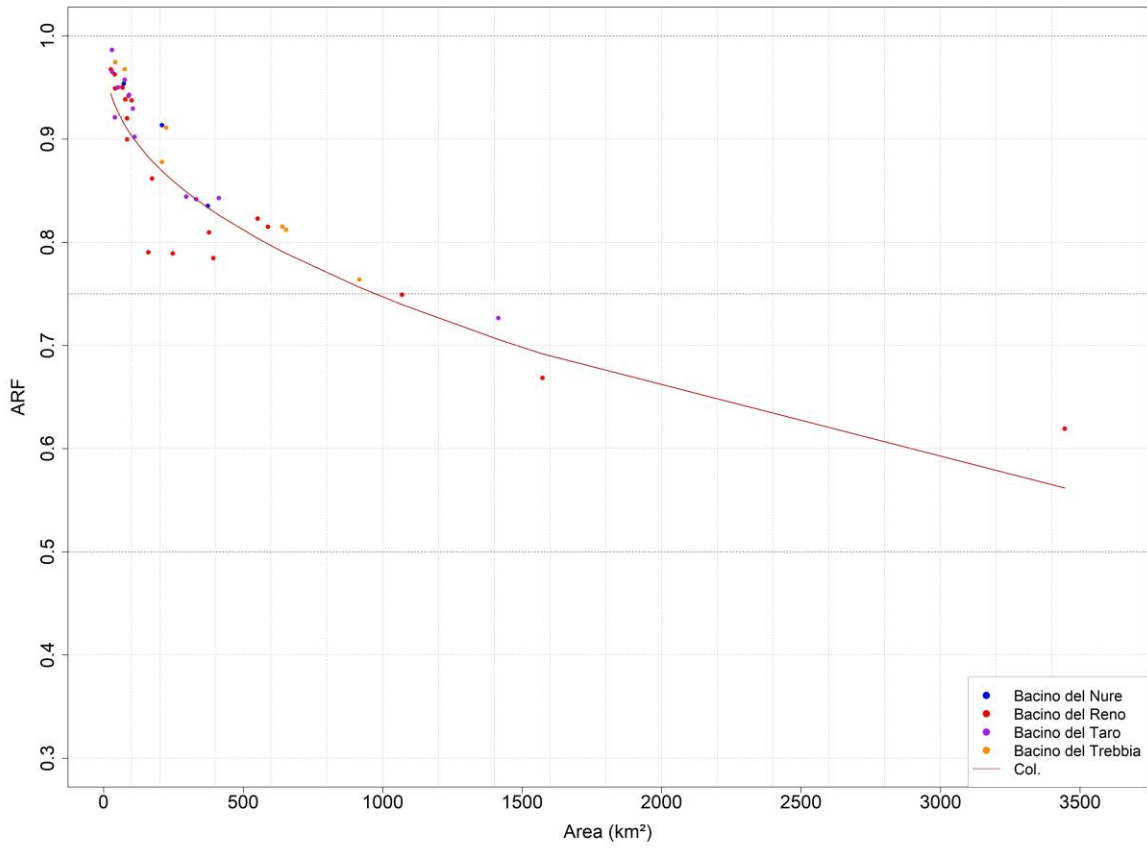
Mentre quelli ottenuti dall'Autore sono:

Parametri	Columbo
c ₁	0,06
c ₂	0,4
c ₃	0,003
c ₄	0,6

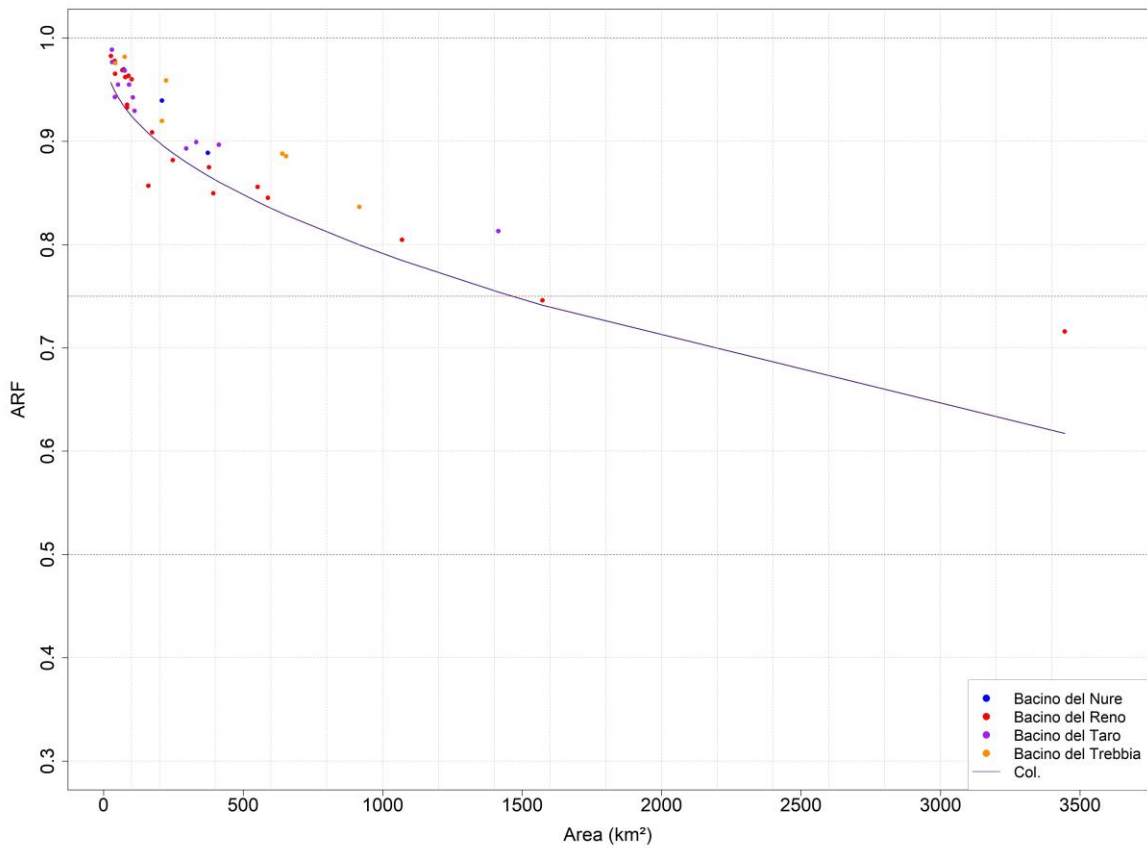
L'espressione di Columbo è stata originariamente formulata per eventi di durata inferiore alle 24 ore e aree al di sotto di 50 km². A partire da questa osservazione, si considerino le curve ricavate per ciascuna durata e i dati sperimentali a partire dalle quali sono state ottenute in Figura 4.10 a-g:



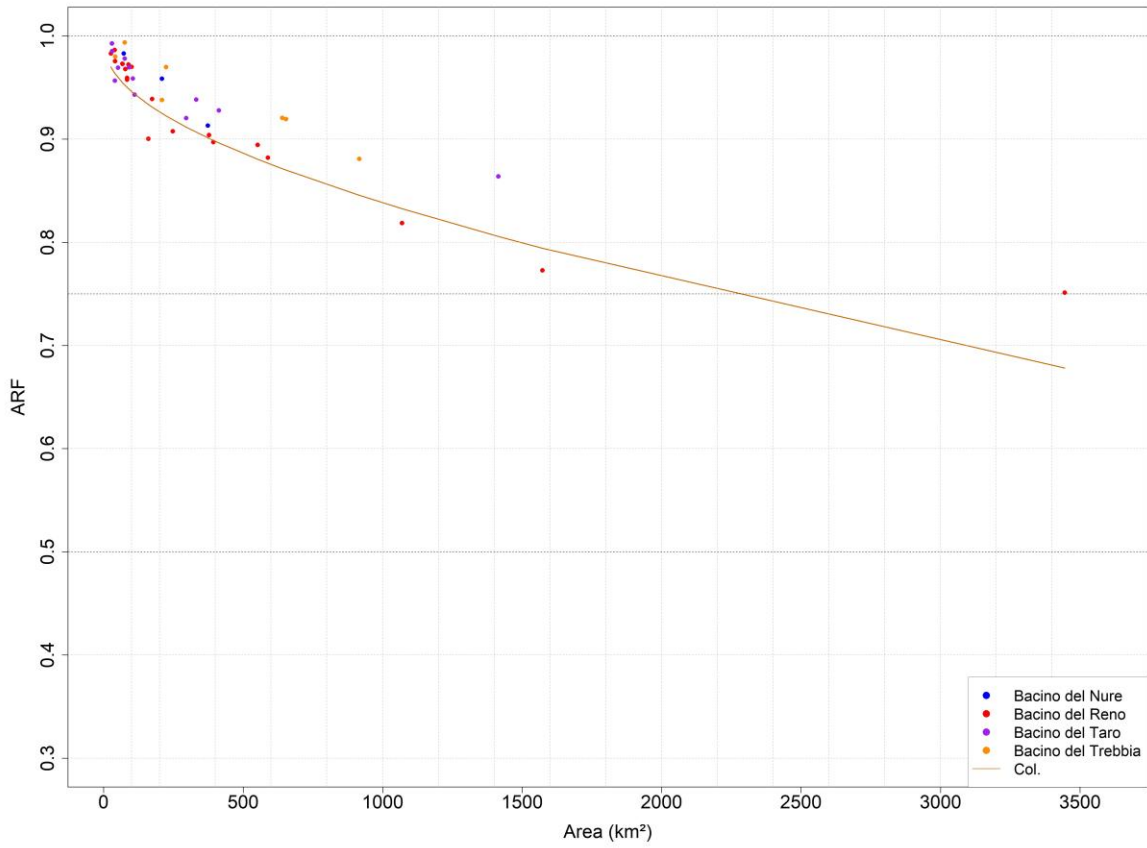
Curva ARF(A, d) ricalibrata per una durata di 3 h (Columbo)



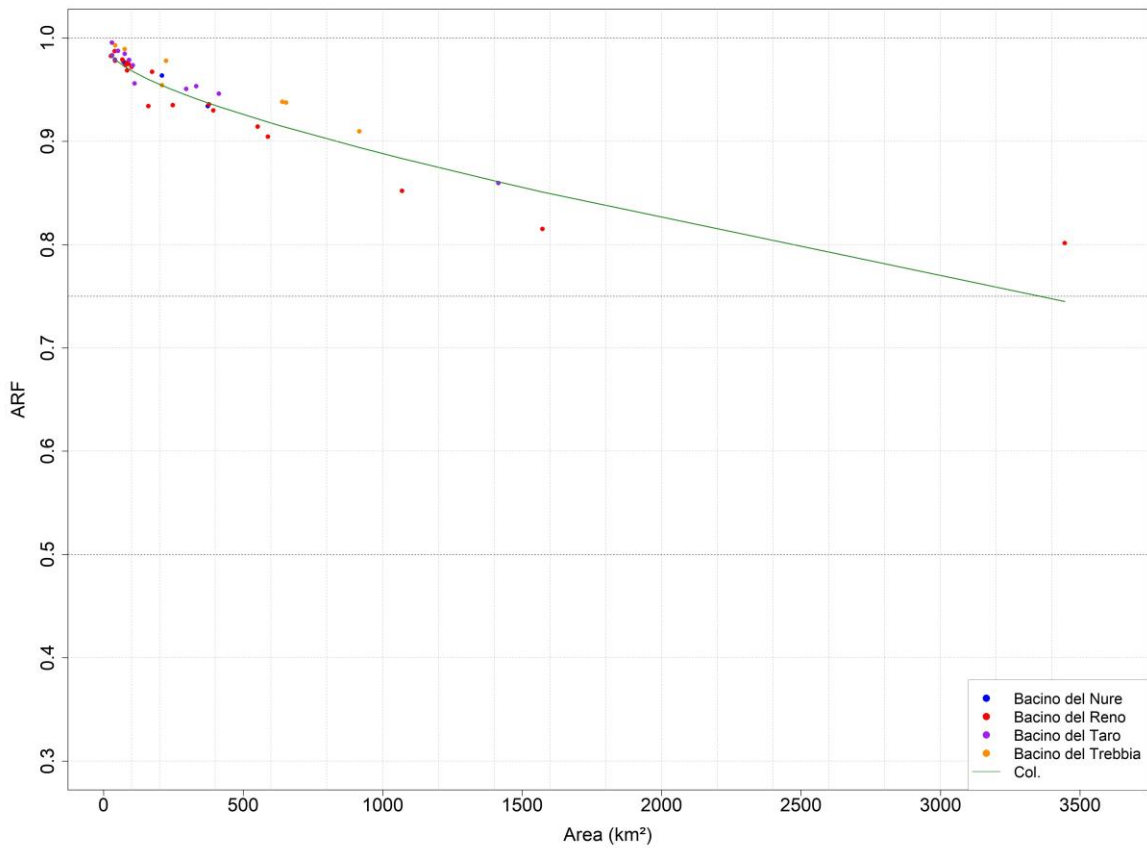
Curva ARF(A, d) ricalibrata per una durata di 6 h (Columbo)



Curva ARF(A, d) ricalibrata per una durata di 12 h (Columbo)



Curva ARF(A, d) ricalibrata per una durata di 24 h (Columbo)



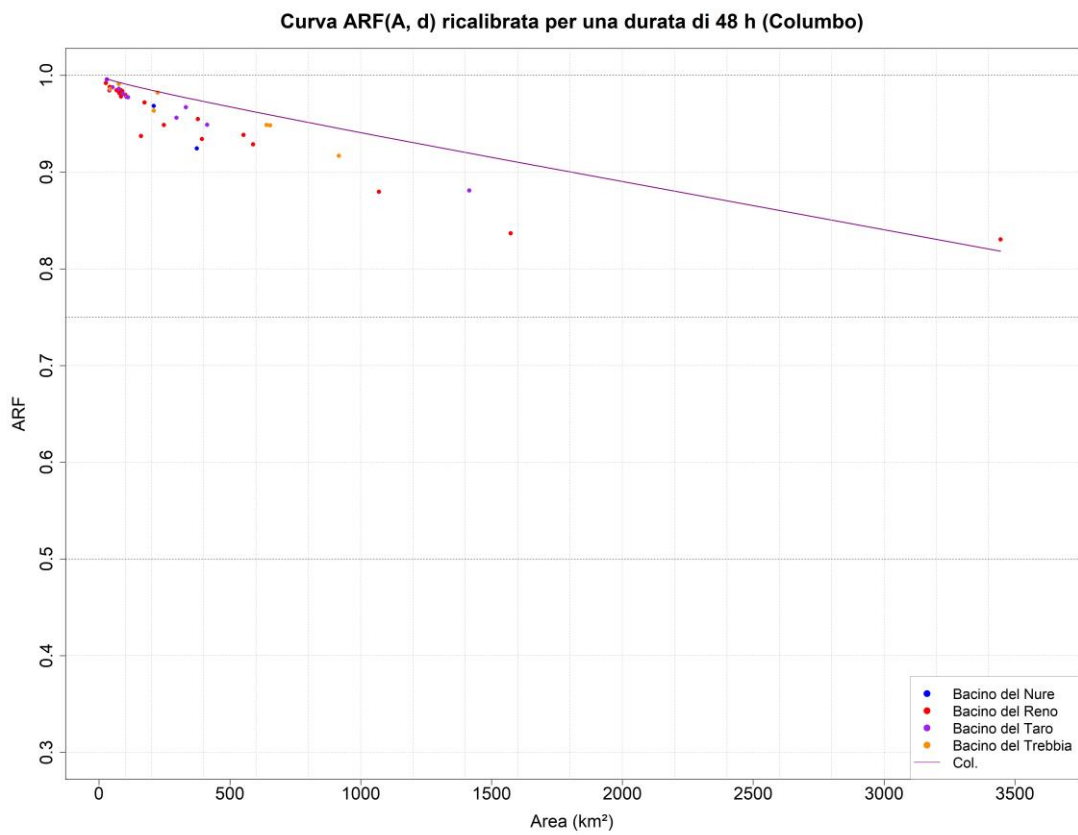
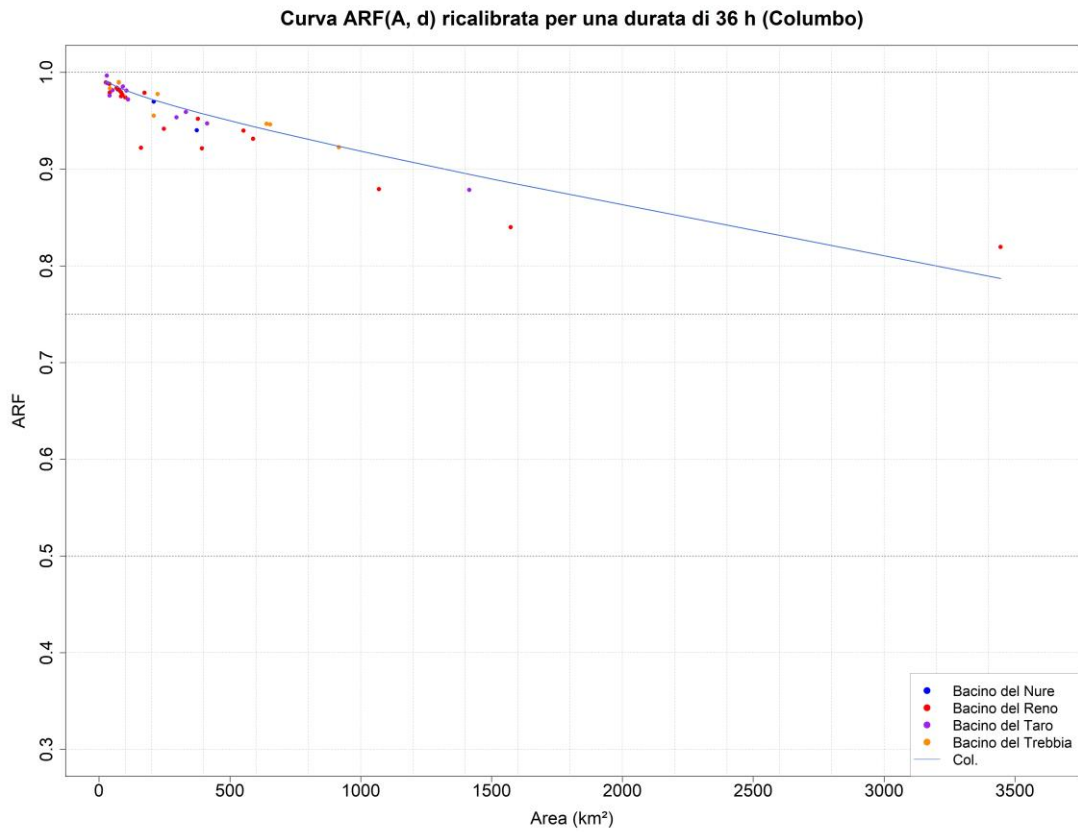


Figura 4.10 Curve ARF(A, d) descritte dalla formulazione proposta da Columbo ricalibrate su valori sperimentali ricavati sui bacini appenninici esaminati, fissata la durata di 1 (a), 3 (b), 6 (c), 12 (d), 24 (e), 36 (f), 48 (g).

Bisogna infatti considerare che le metriche di prestazione ad esse relative:

Columbo	
SSQ	0,3204
NSE	0,8684
NSE _{adj}	0,8670
R ²	0,8695
R ² _{adj}	0,8682
BIAS _{rel}	0,000785

Si tratta di valori soddisfacenti e prossimi a quelli ottenuti dalla USWB; d'altra parte, per ogni durata, invece, si ha:

Durata (h)	NSE _{adj}
1	0,7780
3	0,8410
6	0,6876
12	0,7365
24	0,8674
36	0,8133
48	0,5864

Le prestazioni delle curve valutate fissando la durata, e non prendendo l'intero campione, sono altalenanti, e per 48 h sono addirittura non accettabili.

La proposta di letteratura, se confrontata con il risultato della taratura:

$$ARF(A, d) = (1 - 0,021A^{0,39})d^{(0,0054A^{0,39})}$$

È tale che qualunque sia il dominio considerato, la curva ottenuta dallo studio di Columbo è inadatta a rappresentare i risultati sperimentali ottenuti (Figura 4.11).

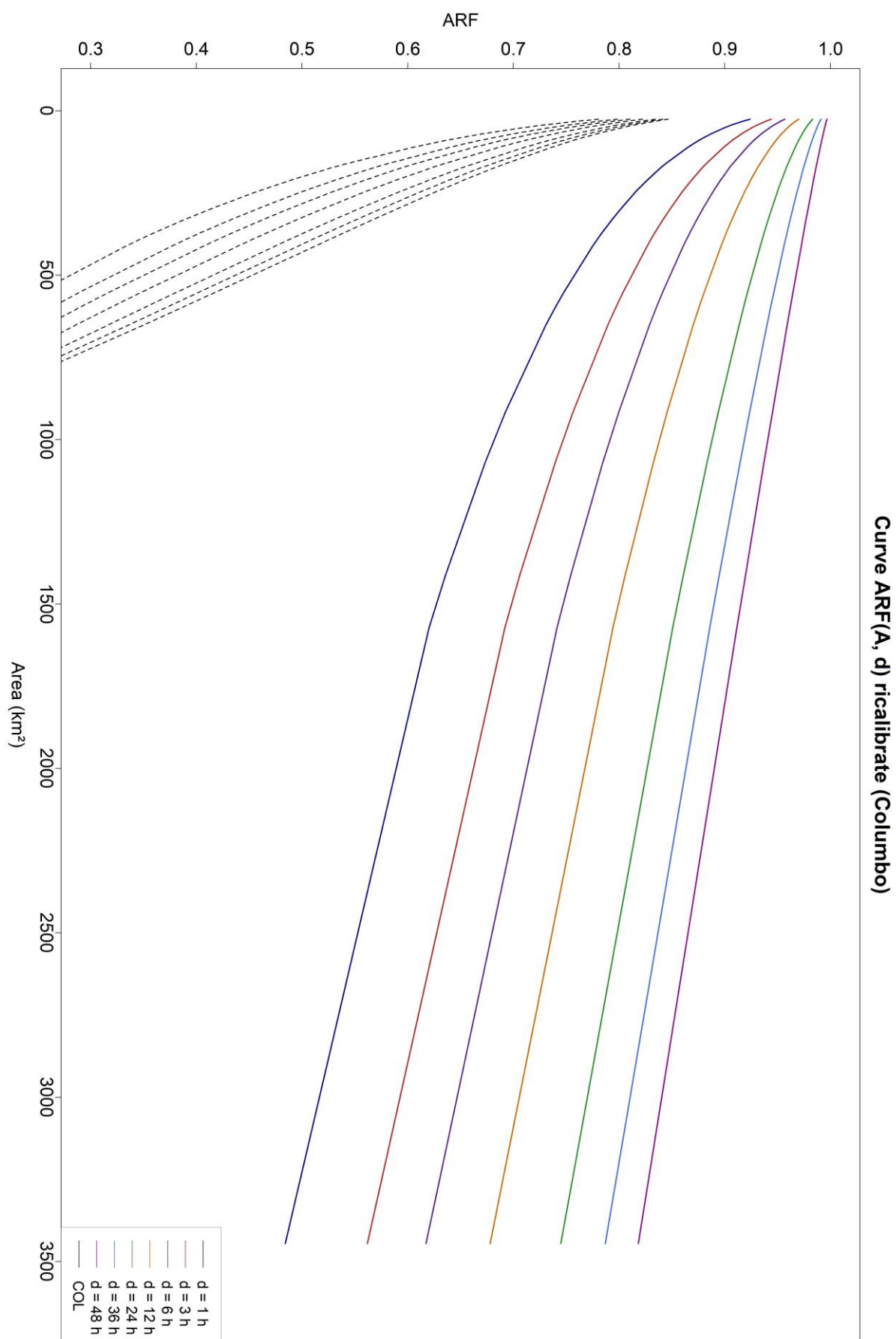


Figura 4.11 Confronto con la curva ottenuta da Columbo.

4.2 Conclusioni

Si riporta una tabella riassuntiva delle metriche ottenute per ciascun modello empirico adattato ai dati sperimentali ricavati, secondo la procedura descritta nel Capitolo 2.

Metriche	Columbo	USWB	M. & P.
SSQ	0,3204	0,3405	0,1973
NSE	0,8684	0,8601	0,9189
NSE _{adj}	0,8670	0,8591	0,9178
R ²	0,8695	0,8630	0,9191
R ² _{adj}	0,8682	0,8620	0,9180
BIAS _{rel}	0,000785	0,00602	0,00229

In generale, si osserva che le prestazioni sono buone per l'USWB e Columbo e ottime per la formulazione di Moisello e Papiri. Dall'analisi risulta quindi che la legge più adatta per rappresentare i dati sperimentali ottenuti è la seguente:

$$ARF(A, d) = 1 - \exp \left[-7,88A^{-0,3} d^{(0,26 - \exp(-1,01A^{0,22}))} \right]$$

Bisogna però tenere presente che i modelli calibrati integralmente sui dati sperimentali valutati hanno prestazioni differenti sulla base della durata considerata, e dei limiti di applicazione con cui queste relazioni empiriche sono state concepite, per non parlare del diverso contesto geografico su cui sono state elaborate. Ci si aspettava infatti che la curva dell'U.S. Weather Bureau, sebbene fornisca una buona performance, risultasse la più idonea, sulla base di studi precedenti sul solo bacino del Reno; d'altra parte, in questa sede si sono presi bacini principalmente montano, in un contesto climatico ben distinto da quello statunitense. Inoltre, il campione sperimentale si costituisce di un'elevata densità per aree ridotte e intermedie e una scarsa rappresentanza di aree di maggiore estensione, fattore che potrebbe aver influenzato le metriche globali calcolate.

Capitolo 5: Valutazione della dipendenza del coefficiente di ragguglio delle piogge all'area dal tempo di ritorno con riferimento agli eventi alluvionali del maggio 2023 in Emilia-Romagna

5.1 Gli eventi alluvionali del maggio 2023 in Emilia-Romagna

Nel maggio 2023, si è consolidato un significativo cambiamento nella configurazione meteorologica dell'Europa. Questo periodo è stato caratterizzato dalla persistenza di anticloni estesi dal Nord Atlantico alla Penisola Iberica e fino alla Scandinavia, mentre si sono verificate depressioni atmosferiche sull'Italia e nel Mediterraneo centrale.

Da un lato, questa configurazione atmosferica ha portato a una frequente e abbondante precipitazione su gran parte del territorio italiano, contribuendo a mitigare la prolungata siccità del 2022; d'altro canto, queste condizioni metereologiche hanno portato a episodi alluvionali, con particolare impatto sull'Emilia orientale e sulla Romagna, a seguito delle straordinarie precipitazioni verificatesi tra l'1-3 e il 16-17 maggio 2023.

In Emilia-Romagna, le precipitazioni ammontano, per l'intero periodo compreso tra i due eventi, un totale medio di 250,7 mm, superando del 230,2% il valore medio climatico massimo storico del 1961. I dati mostrano un evento eccezionale, con anomalie pluviometriche particolarmente intense nelle zone collinari e montuose orientali, dove si sono verificati picchi fino al +500% rispetto alla media climatica del periodo 2001-2020. Nella parte occidentale della regione, le anomalie, sebbene presenti e positive, sono significativamente inferiori, con una media di circa +50%.

Gli eventi estremi hanno causato danni considerevoli su vaste aree del territorio, anche a causa delle difficoltà incontrate dalle acque nel defluire verso il mare. Si sono verificate 23 esondazioni, rotture degli argini, allagamenti di centri urbani e campi coltivati, con danni gravi, oltre a centinaia di frane e più di 700 strade interrotte. Sono stati coinvolti 41 comuni e migliaia di persone sono state evacuate. Complessivamente, nei due eventi climatici estremi, 17 persone hanno perso la vita, di cui 2 nel primo evento e 15 nel secondo (una delle quali a causa di un malore, mentre le altre sono considerate vittime dirette).

5.1.1 Gli eventi metereologici dell'1-4 maggio

Tra il 1° e il 3 maggio, la perdita di quota di una saccatura di origine polare², ha dato origine a una vasta area di bassa pressione nel Mediterraneo centrale, con la formazione di un minimo di pressione nei bassi strati sopra il Tirreno centrale. La presenza di un promontorio sull'Europa orientale ha ostacolato lo spostamento verso est di questo minimo, che ha quindi mantenuto una posizione quasi stazionaria con centro di rotazione sopra l'Italia centrale per l'intera giornata del 2 maggio. Nel corso della giornata del 3 maggio, la saccatura si è gradualmente spostata verso est, mentre il promontorio sull'Europa occidentale si è esteso fino alle regioni settentrionali italiane, portando condizioni meteorologiche in miglioramento nella regione.

La circolazione ciclonica, associata al minimo di pressione posizionatosi sull'Italia centrale, ha facilitato l'incontro tra una massa d'aria umida proveniente dal sud-est, che si stava dirigendo verso il settore adriatico, e correnti d'aria più fredde provenienti dal nord-est. Da qui nascono le continue precipitazioni che hanno colpito la regione, con accumuli fino a oltre 250 mm. L'influenza della catena appenninica ha amplificato il fenomeno sul versante sopravvento, con significative altezze di precipitazione di oltre 200 mm nella zona collinare tra il Modenese e la Romagna.

Le piogge, quindi, hanno interessato principalmente le aree pedecollinari e montuose della parte centro-orientale dell'Emilia-Romagna. La quasi immobilità della depressione atmosferica durante l'intera giornata di martedì 2 maggio e il suo lento spostamento verso est nelle prime ore di mercoledì 3 hanno generato precipitazioni persistenti e di durata prolungata. L'evento si è protratto per circa 48 ore consecutive, producendo quantità così significative di pioggia, in assenza di attività convettive, tipiche delle tempeste estive, ma piuttosto le piogge sono state diffuse e costanti. La tipologia stratiforme dei fenomeni, prevalentemente di debole e moderata intensità (intensità medie orarie attorno ai 2-4 mm/h), ha fatto registrare bassi valori di cumulata sull'ora durante tutto l'evento, a differenza delle cumulate giornaliere eccezionali osservate nel Bolognese ed in Romagna.

Nella Figura 5.1, è rappresentata la precipitazione aggregata sull'intero periodo dell'evento.

² Una saccatura di origine polare è un concetto meteorologico che si riferisce a una regione di bassa pressione atmosferica che si forma o si sposta dalle regioni polari o subpolari verso le latitudini più basse, come quelle temperate o tropicali. Questa bassa pressione atmosferica è associata a un'area di instabilità atmosferica in cui l'aria più fredda proveniente dalle regioni polari incontra l'aria più calda delle regioni più basse.

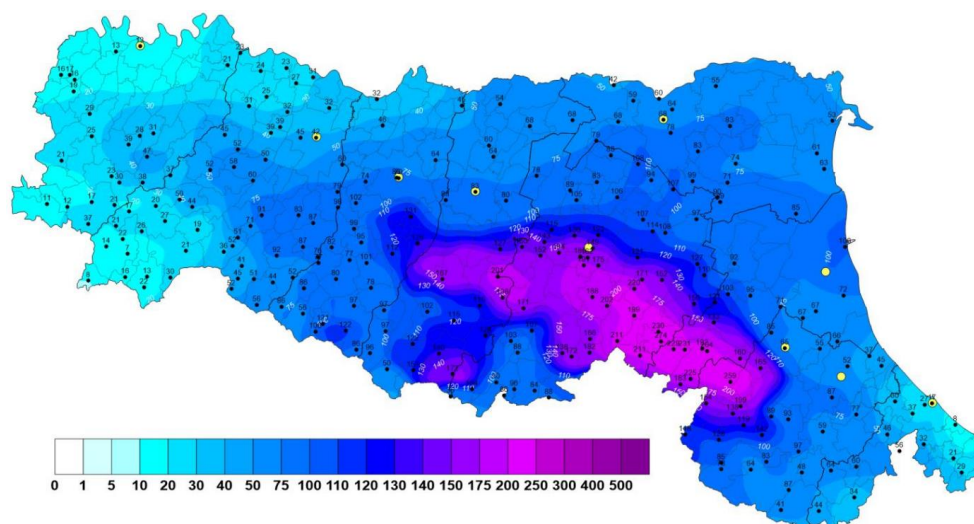


Figura 5.1 Riepilogo delle precipitazioni cumulate durante l'evento dal 1 al 3 maggio 2023; i maggiori accumuli, spesso superiori a 200 mm, si sono concentrati principalmente nel settore centro-orientale della regione. (Struttura Idro-Meteo-Clima, Arpae).

La straordinarietà dell'evento non sta nelle intensità delle precipitazioni, inferiori ai massimi registrati nella regione negli anni recenti, ma nella sua durata e nella vasta estensione territoriale coinvolta. Confrontando i dati cumulati delle precipitazioni sull'intero territorio regionale, si è rivelato il più intenso evento pluviometrico per 48 ore consecutive dal 1997 e il più significativo nella stagione primaverile dal 1961.

L'evento di precipitazione che si è verificato dall'1 al 3 maggio è stato preceduto da un mese di aprile caratterizzato da precipitazioni di modesta entità, concentrate principalmente nel settore centro-orientale della regione. Nel frattempo, il resto della regione aveva registrato una notevole siccità sin dall'inizio dell'anno idrologico. Sebbene il terreno fosse inizialmente in uno stato di deficit idrico, particolarmente evidente nella zona bolognese, che ne aveva ridotto la capacità di assorbire le precipitazioni, queste ultime hanno rapidamente saturato lo strato superficiale del suolo, generando ingenti quantità di acqua che hanno causato significativi aumenti di flusso lungo tutti i corsi d'acqua del settore centro-orientale della regione, con eventi di piena su svariati bacini collinari, piuttosto che nelle zone montuose con rotte degli argini e esondazioni, oltre a centinaia di fenomeni franosi, che spaziavano da piccoli smottamenti a frane di considerevoli dimensioni.

La natura stratiforme dell'evento, con precipitazioni di intensità non elevata, che però hanno stazionato in modo prolungato sul territorio, è ben evidente dalla registrazione del pluviometro a Fontanelice, una delle località più colpite del bolognese, che ha raggiunto quasi 280 mm.



Figura 5.2 Dati registrati dal pluviometro a Fontanelice (dalle 00UTC del 1° maggio alle 12:00UTC del 3 maggio), (Fondazione CIMB).

Le misure orarie (istogramma rosso) si aggirano intorno a circa 10 mm/h per quasi tutta la giornata del 2 maggio: la persistenza nello spazio e nel tempo ha dato luogo ad una risposta elevata dei bacini in termini di portata.

Piene rilevanti si sono verificate nei bacini di Samoggia, Idice, Quaderna, Sillaro, Santerno, Senio, Lamone e Montone, contemporaneamente interessati da piene superiori ai massimi storici registrati.

5.1.2 Gli eventi meteorologici dell'16-18 maggio

L'evento dei primi di maggio è stato seguito da precipitazioni in alcuni casi anche abbondanti, che però non sono paragonabili a quelle dell'inizio del mese; nonostante ciò, hanno avuto particolare impatto in alcuni frangenti per via dei suoli già saturi. In particolare, durante la giornata del 13 maggio, le precipitazioni si sono manifestate principalmente sotto forma di temporali convettivi e non hanno portato ad accumuli significativi nell'arco delle 24 ore.

Queste precipitazioni hanno causato vari disagi nella provincia di Bologna, come riportato dai resoconti locali. Nel centro della città, le piogge intense hanno portato all'esondazione del torrente Ravone, causando l'allagamento di via Saffi e rendendo necessario l'intervento dei vigili del fuoco. Altri casi di allagamenti sono stati segnalati anche su via Saragozza e via di Casaglia. Inoltre, in alcune zone del bolognese, i temporali hanno causato ulteriori movimenti franosi, aggravando una situazione già critica a causa delle precipitazioni dei giorni precedenti.

L'evento più significativo è stato quello tra il 16 e il 17 maggio, quando ingenti precipitazioni, simili per quantità e aree interessate all'evento dell'1-3 maggio, hanno colpito nuovamente la

parte centro-orientale della regione, che aveva già subito un grave evento meteorologico, con una notevole incidenza sui rilievi da Bologna a Forlì-Cesena.

A partire dal 16 maggio, si sono osservati repentini aumenti dei livelli idrometrici su tutti i corsi d'acqua, sia quelli principali che quelli minori, nel settore centro-orientale della regione. Questi aumenti dei livelli idrometrici sono stati più pronunciati nei tratti montani dei fiumi e si sono accumulati nei tratti vallivi, dando origine a onde di piena con volumi molto elevati. Su fiumi come il Samoggia, l'Idice, il Quaderna, il Sillaro, il Santerno, il Senio, il Lamone, il Montone, il Ronco, il Bevano, il Savio e il Rubicone, sono stati raggiunti livelli di piena ancora più elevati rispetto a quelli osservati durante l'evento di 14 giorni prima. In alcuni punti, i livelli idrometrici hanno superato persino i massimi livelli registrabili dagli strumenti di monitoraggio. La straordinaria altezza e durata delle piene ha causato danni considerevoli sul territorio, che hanno superato quelli provocati dalla piena iniziale di maggio. Questi danni sono stati evidenti sia nei tratti vallivi che in quelli pedecollinari dei corsi d'acqua, con fenomeni di esondazione, erosione intensa, inondazioni e trasporto di detriti, oltre a superamenti degli argini; ma anche l'occlusione delle arcate dei ponti con conseguente rigurgito delle acque a monte, nuove rotte e l'aggravamento di quelle precedentemente ripristinate con misure provvisorie nei giorni precedenti.

La precipitazione media regionale durante le 48 ore è stata di 73,2 mm, e il giorno 16 sono stati registrati massimi giornalieri eccezionali, tra cui fino a 200 mm. Alla fine di questo evento, alcune stazioni meteorologiche avevano accumulato oltre 600 mm di pioggia dall'inizio del mese. In altre parole, le precipitazioni avvenute nel periodo tra il 16 e il 17 maggio più intense e diffuse rispetto a quelle dell'inizio del mese, hanno portato a cumulate di precipitazione che hanno superato i massimi storici registrati in 17 giorni sul territorio regionale dal 1961.

Dal punto di vista meteorologico, nella giornata del 15 maggio si osservano due elementi principali: a ovest, è presente un'area di alta pressione tra l'Atlantico e la penisola Iberica; ad est, si estende un ampio promontorio dal Mediterraneo orientale fino all'Europa nord-orientale. Tra queste due configurazioni bariche, si forma una depressione atmosferica che alimenta un vortice posizionato tra la Tunisia e la Sicilia. Questo vortice depressionario, chiamato “Minerva” dal Servizio Meteorologico dell'Aeronautica Militare, è stato il principale motore delle condizioni meteorologiche instabili che hanno interessato prima le regioni meridionali della penisola, poi quelle centrali e infine l'Emilia-Romagna tra il 15 e il 17 maggio 2023. All'inizio del 16 maggio, Minerva, spostandosi verso le regioni centrali

dell'Italia, ha generato una circolazione ciclonica intensa. Questo ha portato correnti umide di Scirocco dal Mediterraneo sud-orientale verso l'area adriatica, mentre forti venti di Bora hanno colpito le coste dell'Adriatico settentrionale. L'interazione tra queste masse d'aria e la convergenza dei venti a bassa quota, soprattutto nelle vicinanze della Romagna, ha scatenato intense precipitazioni, comprese alcune di natura convettiva, durante le prime ore del mattino. Questa è l'unica fase dell'evento caratterizzata anche da fenomeni convettivi, innestati in un sistema stratiforme.

Nel corso dell'intera giornata di martedì 16 e la mattina di mercoledì 17 maggio, il vortice depressionario, con il suo centro praticamente fermo sull'Italia centrale, ha continuato a trasportare masse d'aria ricche di umidità verso la pianura padana, provocando precipitazioni persistenti, soprattutto nel settore centro-orientale dell'Emilia-Romagna. L'effetto di ostacolo creato dalla catena appenninica, simile a quanto accaduto durante l'evento del 2-3 maggio, ha concentrato e intensificato le precipitazioni sul versante a monte dei rilievi (effetto Stau), in particolare nelle zone collinari tra il Bolognese e la Romagna. Sebbene nella seconda metà della giornata di mercoledì 17 maggio il vortice depressionario inizi a indebolirsi, continuano a mantenersi condizioni di instabilità, seppur meno intense, per il resto della giornata.

Le altezze di pioggia osservata durante l'evento meteorologico che ha interessato tutto il territorio regionale il 16 maggio, sono ampiamente superiori ai 100 mm. Le province più colpite sono state Forlì-Cesena, Ravenna, Bologna e Rimini, con massimi di precipitazione giornaliera il 16 maggio di fino a 200 mm.

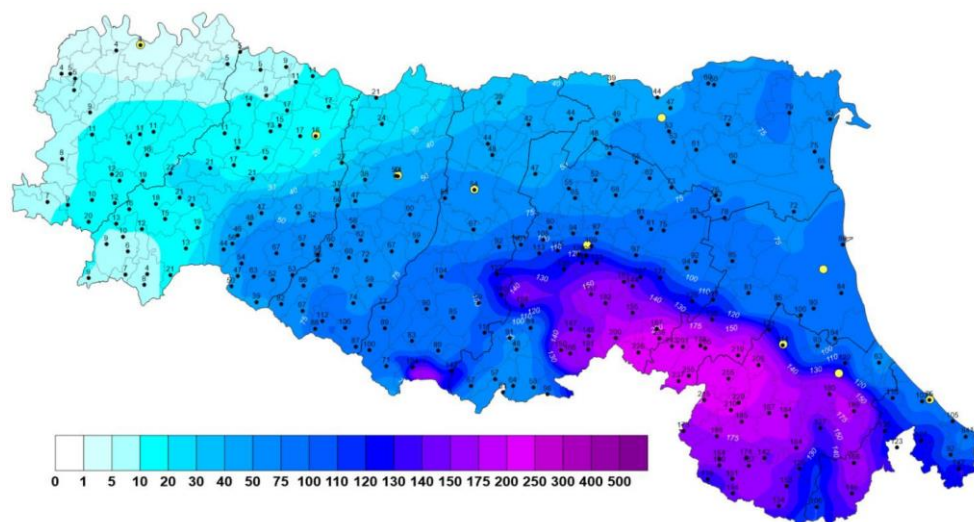


Figura 5.3 Precipitazioni cumulate durante l'evento che ha interessato i giorni 16 e 17 maggio 2023; le aree più colpite, con valori superiori a 200 mm, si concentrano nella zona appenninica centro-orientale della regione, compresi i territori tra il Bolognese e il Riminese, dove si sono registrati massimi di precipitazione su due giorni consecutivi fino a 260 mm (Struttura Idro-Meteo-Clima, Arpae).

Dalla Figura 5.3 si osservano le precipitazioni cumulate durante l'intero evento che ha interessato l'Emilia-Romagna ottenute attraverso il dataset ERG5.

Questo evento meteorologico, complessivamente, è stato eccezionale sia per l'intensità delle precipitazioni che per l'ampia estensione territoriale coinvolta. L'analisi statistica delle precipitazioni massime registrate nelle diverse stazioni, con diverse durate (1, 3, 6, 12, 24 e 48 ore), svolta dalla Struttura IdroMeteoClima, ha rivelato tempi di ritorno superiori ai 200 anni per alcune stazioni pluviometriche nei bacini di Santerno, Senio, Lamone, Montone e Ronco, per le piogge di durata 24 ore. Per le piogge di durata 2 giorni consecutivi, si sono stimati tempi di ritorno superiori ai 200 anni. Di seguito, si riporta la registrazione del pluviometro a Casola Valsenio, uno dei comuni più colpiti:

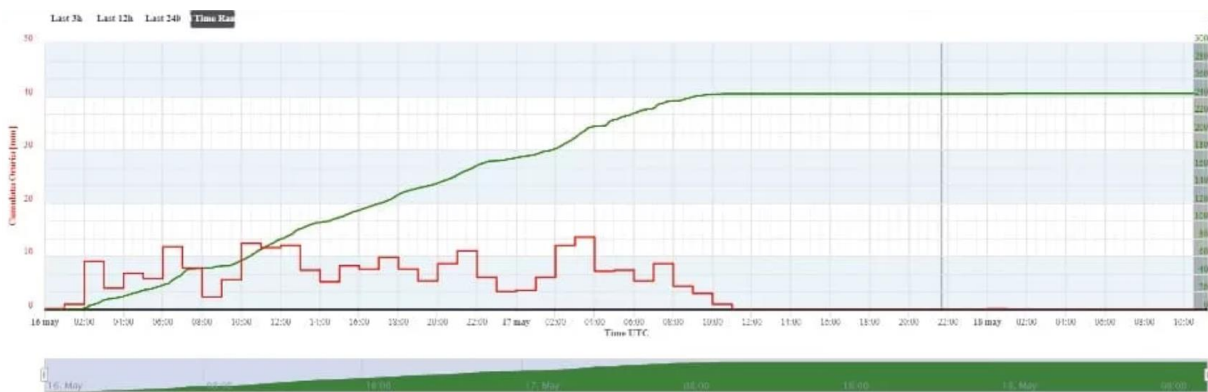


Figura 5.4 Registrazione delle altezze di pioggia a passo orario e cumulate durante l'evento tra dalle 00UTC del 16 maggio alle 00UTC del 18 maggio (Fondazione CIMA).

Nuovamente, è evidente la natura stratiforme dell'evento, con precipitazioni ad intensità non elevata ma molto persistenti nello spazio e nel tempo.

Prendendo in esame l'intero periodo dal 1° al 17 maggio, l'evento di precipitazione avvenuto corrisponde al nuovo record storico della cumulata di precipitazione in 17 giorni in molte stazioni pluviometriche, coprendo un'area che va dal bacino del Samoggia a quello del Ronco. In queste zone, i valori medi della cumulata si sono attestati sopra i 300-400 mm, raggiungendo massimi di circa 600 mm. Con riferimento all'intera regione, la media areale per l'intero evento è stata di 221,40 mm, superiore rispetto a qualsiasi altro massimo storico per la medesima durata dal 1961.

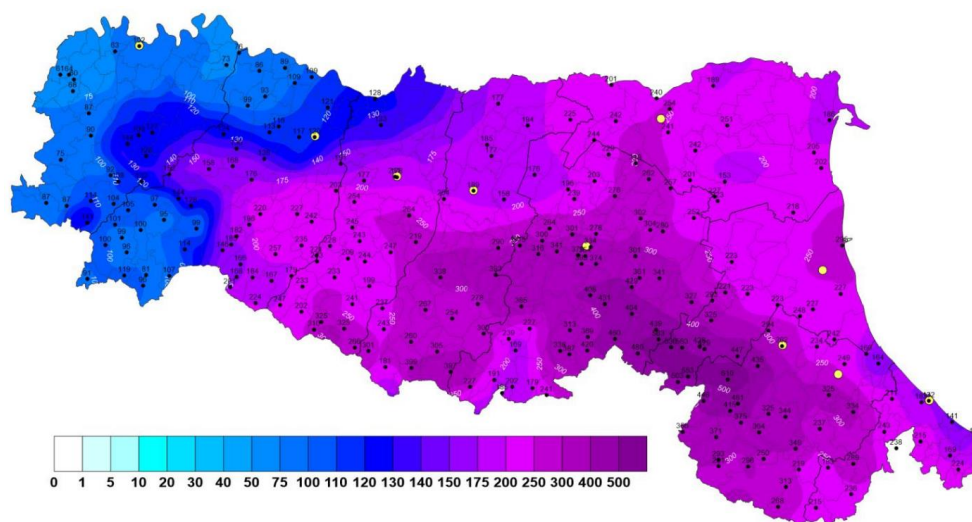


Figura 5.5 Precipitazione cumulata sull'intero periodo, 1°-17 maggio, con indicazione dei valori puntuali e dei confini dei territori comunali (Struttura IdroMeteoClima, Arpae).

In dettaglio, considerando l'area appenninica bolognese e romagnola, rispetto al periodo dal 1961 al 2020, entrambi gli eventi pluviometrici dell'1-3 maggio e del 16-17 maggio 2023 superano ampiamente la precipitazione media mensile e rappresentano insieme circa il 25% del totale annuo. Nel complesso, il totale delle precipitazioni di maggio 2023 corrisponde a quasi il 60% dell'accumulo climatico annuale.

L'evento del maggio 2023 rappresenta quindi il record assoluto per l'accumulo di precipitazione in 24 e 48 ore consecutive e anche nel periodo di 17 giorni in esame, superando il precedente massimo storico.

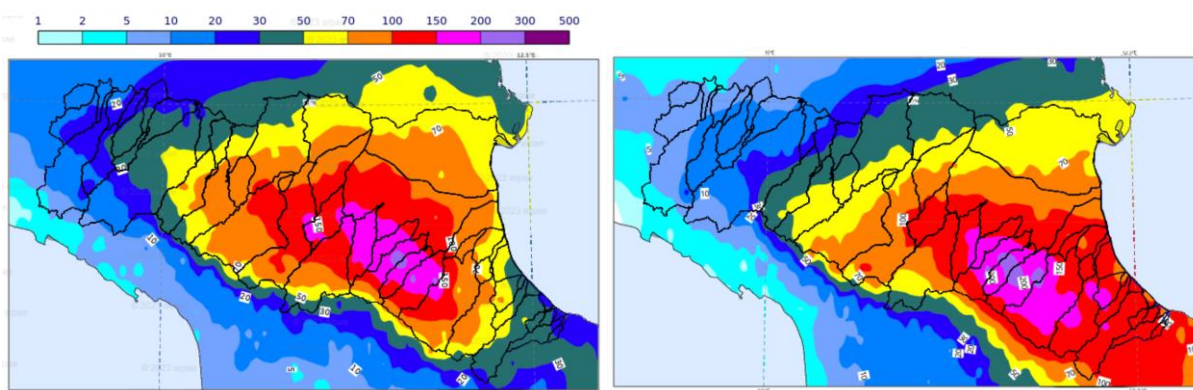


Figura 5.6 Mappa a sinistra: precipitazione cumulata dell'evento dalle ore 12 del 1° maggio alle ore 12 del 2° maggio sui bacini della regione Emilia-Romagna. Mappa a destra: precipitazione cumulata dell'evento dalle ore 00 del 16 maggio alle ore 00 del 17 maggio sui bacini idrografici della regione Emilia-Romagna (Struttura IdroMeteoClima, Arpae).

Le precipitazioni hanno interessato una vasta area del territorio regionale durante entrambi gli eventi. I bacini che hanno ricevuto le precipitazioni più intense, dal Sillaro al Ronco, coincidono con quelli colpiti dalle forti precipitazioni e dalle piene dell'evento dell'1-3

maggio. Tuttavia, l'evento di precipitazione del 16-17 maggio si è caratterizzato non solo per la sua maggiore intensità e ampiezza, ma anche per le condizioni iniziali di saturazione del suolo, se non addirittura di totale saturazione, dovute agli eventi precedenti, e quindi ancora una volta le forti precipitazioni sono infiltrate solo parzialmente e sono andate a ingrossare i corsi d'acqua.

In conclusione, le alluvioni consecutive che hanno colpito l'Emilia-Romagna rappresentano due eventi meteorologici estremi, che rientrano in una tendenza climatica già evidenziata da decenni di dati. Le precipitazioni estreme verificatesi sono il risultato di una particolare configurazione meteorologica atmosferica che è rimasta stabile sull'Europa, che, a causa di un'alta pressione stabile a ovest e un'alta pressione a est, ha contribuito all'istaurarsi di un periodo prolungato di condizioni meteorologiche perturbate nel Mediterraneo centrale. In entrambi gli scenari meteorologici, sia quello del 1° maggio che del 16 maggio, due impulsi provenienti dal Nord Africa hanno causato la formazione di un centro di bassa pressione (ciclone) sull'Italia, portando a precipitazioni abbondanti nelle regioni dell'Emilia-Romagna e delle Marche, così come lungo il medio-basso Tirreno. L'effetto combinato dell'effetto Stau, che si verifica quando le correnti atmosferiche colpiscono le catene montuose e causano il raffreddamento e la condensazione dell'aria, insieme ai venti persistenti che trasportano l'umidità dal mare verso l'entroterra, ha provocato l'accumulo di ingenti quantità d'acqua (ad esempio, 300 millimetri sulle colline del forlivese e 135 millimetri nelle pianure circostanti, in maniera diffusa e non puntuale) su terreni che erano già saturi d'acqua.

5.1.3 Le macro-aree regionali ARPAE

Per coordinare le operazioni di protezione civile, la regione è divisa in otto macroaree, selezionate in base alle caratteristiche climatiche ed idrologiche omogenee, cercando, quando possibile, di seguire i confini amministrativi. Di seguito è elencato l'ordine delle macroaree da est a ovest:

- A - Bacini romagnoli (RA, FC, RN);
- B - Pianura e costa romagnola (RA, FC, RN);
- C - Bacini emiliani orientali (BO, RA);
- D - Pianura emiliana orientale e costa ferrarese (FE, RA, BO);
- E - Bacini emiliani centrali (MO, RE, PR);
- F - Pianura emiliana centrale (MO, RE, PR, BO);

- G - Bacini emiliani occidentali (PR, PC);
- H - Pianura e bassa collina emiliana occidentale (PR, PC).

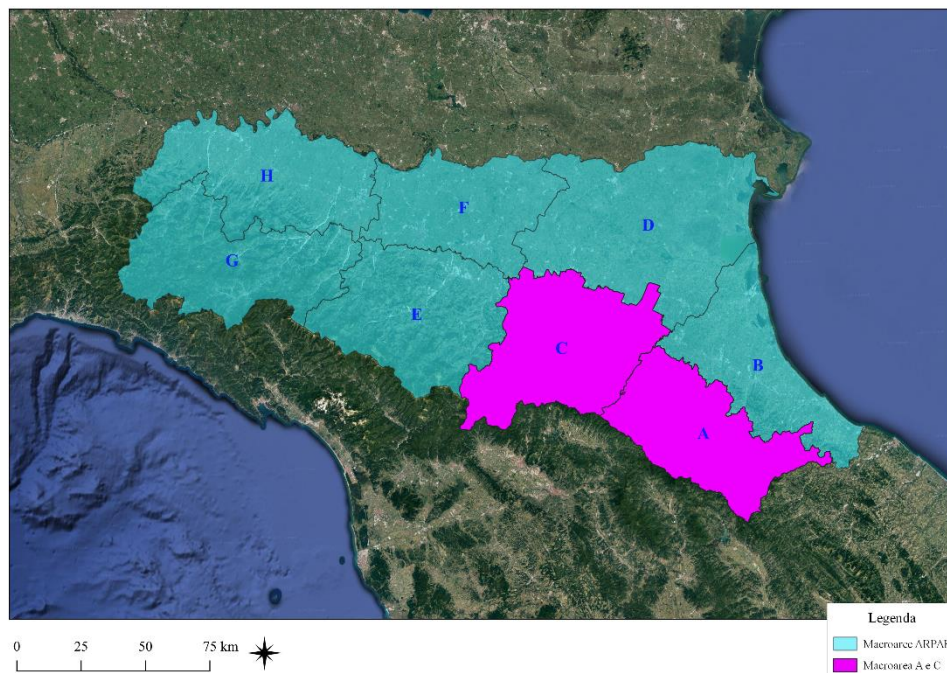


Figura 5.7 Mappa delle 8 Macroaree di allertamento idrologico della Regione Emilia-Romagna, con evidenza delle macroaree oggetto di studio.

Caso di studio del presente capitolo è la valutazione dell'ARF medio annuale, in funzione dell'area e della durata, per le macroaree ARPAE A (2560,7 km²) e C (3083,8 km²), maggiormente colpite dagli eventi eccezionali finora descritti. Per queste zone, si è ritenuto di particolare rilievo il confronto con la valutazione del fattore di riduzione della precipitazione areale medio riferito al 2023 e la distribuzione dell'ARF_Y(A, d) nei 32 anni precedenti.

5.2 Confronto: ARF(A, d) annuale del 2023

Impiegando la medesima procedura sperimentale adottata per la stima dell'ARF_Y(A, d) dei bacini sub-appenninici, è stato valutato il fattore di riduzione della precipitazione all'area per le macro-aree selezionate. Nel dettaglio, è stato considerato, come periodo di riferimento, dapprima l'arco temporale dal 1991 al 2022, per poi svolgere la stima del coefficiente di ragguglio con riferimento al 2023.

5.2.1 Altezza di precipitazione media areale massima annuale

Il fattore di riduzione delle precipitazioni per la j-esima area è valutato come:

$$ARF(A, d) = \frac{h_{A,max,j}(d)}{h_{max,j,m}(d)}$$

La prima parte di questo capitolo è stata dedicata alla descrizione approfondita degli eventi pluviometrici estremi del mese di marzo, non solo in termini di dinamiche meteorologiche che ne hanno causato l'avvenire, ma con riferimento anche ai valori di precipitazione cumulata registrati, evidenziandone l'eccezionalità riportando le valutazioni del tempo di ritorno svolte dal Servizio IdroMeteoClima ARPAE.

Dunque, si è ritenuto di interesse procedere con la stima del tempo di ritorno delle precipitazioni medie areali massime annuali, per ciascuna durata considerata, per analizzare l'effetto degli eventi pluviometrici alluvionali sul valore spaziale della precipitazione per le due macroaree considerate. In Figura 5.7, sebbene la stima spaziale sia riferita all'intero territorio italiano e per i primi 16 giorni di maggio, è evidente che sia tutt'altro che ordinaria, specialmente per via del carattere di eccezionalità dell'evento non solo dalla sua persistenza e, di conseguenza, intensità, ma anche dal punto di vista della sua estensione spaziale. In particolare, si osserva che nonostante per il 2023 sia stata presa metà del mese, l'acqua registrata dall'insieme dei pluviometri italiani è già superiore a quella raccolta in 60 dei 69 mesi di maggio interi del periodo 1951-2019.

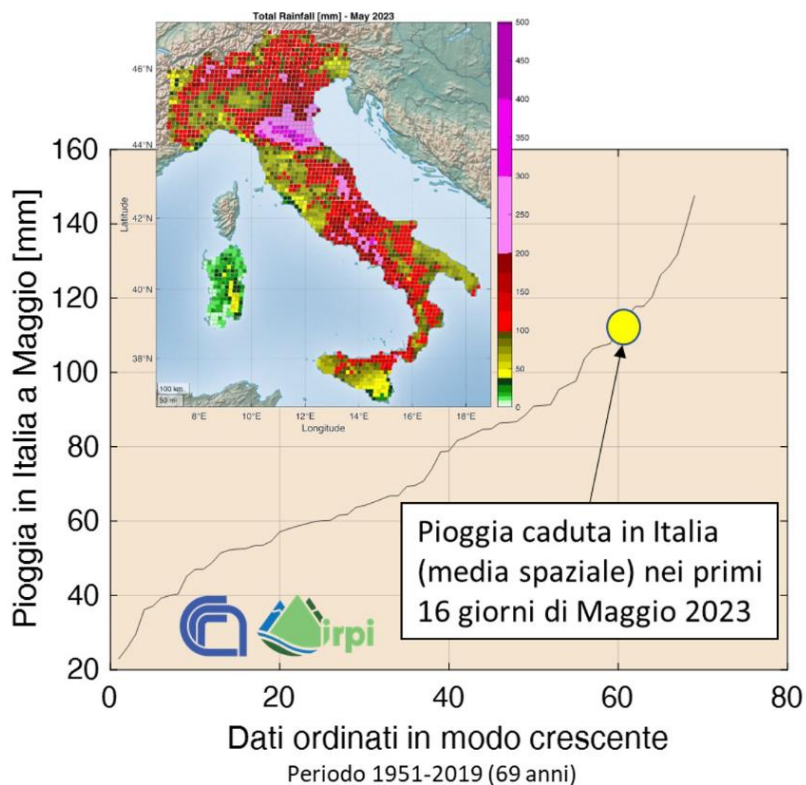


Figura 5.8 Collocazione del caso del 2023 nella distribuzione crescente delle precipitazioni di maggio in Italia (CNS-IRPI).

Il tempo di ritorno T si definisce come il numero medio di anni che intercorre affinché un evento di data magnitudo venga eguagliato o superato. Questo è legato dalla probabilità cumulata di superamento, $P(X > x_T)$ dalla seguente relazione:

$$P(X > x_T) = \frac{1}{T}$$

E di conseguenza alla probabilità di non superamento, $P(X \leq x_T)$ o F_s dalla:

$$P(X \leq x_T) = 1 - P(X > x_T) = 1 - \frac{1}{T}$$

Quindi si ha:

$$T = \frac{1}{1 - F_s}$$

Per il calcolo della cumulata di non superamento empirica, si procede disponendo in ordine crescente i dati, il cui rank quindi andrà da 1 a N , con N numero di osservazioni; dopodiché, si impiega la cosiddetta Weibull plotting position. Questa formula di Weibull permette di calcolare la cumulata di superamento a partire solo dalla posizione occupata dal dato, ovvero:

$$F_s(x_i) = \frac{i}{N + 1}$$

Dove i è la posizione o rank del valore considerato e sarà compresa tra 1 ed N . Questa formula non dipende dal valore del dato, ma dipende solo dalla posizione che occupa nella lista ordinata dal valore minimo al valore massimo.

Dunque finora, senza l'applicazione di alcun modello teorico, si è ricavata a partire dai dati la cumulata di non superamento empirica; d'altra parte, l'impiego di una distribuzione di probabilità teorica spesso è necessaria, ad esempio perché i dati sono rappresentativi solo per determinati anni e si deve estendere questo comportamento per un periodo più lungo, oppure estrapolare delle informazioni. Per questo motivo, ricavata la cumulata di non superamento osservata, si applica una cumulata teorica, ovvero a riprodurre i dati con un modello, come la distribuzione di Gumbel, considerata in questo caso.

La distribuzione di Gumbel è infatti generalmente utilizzata per riprodurre la distribuzione dei valori estremi di una variabile aleatoria, in questo caso la precipitazione massima media areale annuale. L'espressione analitica della distribuzione di Gumbel fa riferimento ad una cumulata di non superamento $P(X < x)$ è definita come:

$$P(X < x) = \exp\left(-\exp\left(\frac{-x - b}{a}\right)\right)$$

Si tratta di una distribuzione doppia esponenziale a due parametri, a e b, che si possono calcolare direttamente dai dati in funzione della media e della varianza:

$$\mu_x = b + 0,577a$$

$$\sigma_x^2 = 1,645a^2$$

L'obiettivo con cui di solito si applica la distribuzione di Gumbel è legato al fatto che permette di calcolare i valori della variabile aleatoria x associati a diversi tempi di ritorno:

$$x = b + a \left[-\ln\left(-\ln\left(1 - \frac{1}{T}\right)\right) \right]$$

Una volta applicato il modello, è necessario valutarne la performance, se riproduce bene il comportamento dei dati o meno.

5.2.1.1 Plotting Position Correlation Coefficient

Per analizzare la bontà del modello, si deve applicare un test statistico. Nell'ambito del lavoro svolto, si è adottato il Plotting Position Correlation Coefficient (PPCC), uno strumento fondamentale nell'analisi climatica e idrologica. Il coefficiente PPCC viene calcolato utilizzando il coefficiente di correlazione di Pearson tra le probabilità cumulative empiriche e quelle previste:

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Dove:

- x_i è l'i-esimo del campione, ordinato in maniera crescente;
- y_i è, nel caso della distribuzione di Gumbel:

$$y_i = -\ln\left(-\ln\left(\frac{i}{N+1}\right)\right)$$

Questo coefficiente misura la forza e la direzione della relazione lineare tra le due serie di probabilità cumulative su una carta probabilistica: un alto valore di PPCC (vicino a 1) suggerisce che il modello o la funzione di distribuzione è adatto ai dati, il che significa che le previsioni basate su quel modello sono in buon accordo con la realtà. D'altra parte, un basso

valore di PPCC (vicino a 0 o negativo) suggerisce che il modello potrebbe non adattarsi bene ai dati e che le previsioni basate su di esso potrebbero non essere affidabili.

5.2.1.2 Test d'ipotesi

Il test d'ipotesi è un procedimento statistico utilizzato per prendere decisioni basate su prove empiriche o dati campionari riguardanti una o più affermazioni, chiamate ipotesi. Il primo passo è formulare due ipotesi: l'ipotesi nulla (H_0) e l'ipotesi alternativa (H_1). L'ipotesi nulla, ad esempio, può essere: il campione è distribuito secondo una distribuzione di Gumbel; l'ipotesi alternativa H_1 in tal caso può essere: il campione non è distribuito secondo la distribuzione di Gumbel. Per stabilire se l'ipotesi nulla sia accettata o meno, si associa all'ipotesi una grandezza rappresentabile tramite variabile casuale di distribuzione nota: se H_0 è vera, la grandezza segue quella determinata distribuzione e assume certe caratteristiche statistiche; se H_0 non è vera, può essere valida la ipotesi alternativa considerata. Definita l'ipotesi nulla come variabile, si stabilisce una regione di accettazione del test, in cui H_0 viene accettata, ed una regione di rigetto, oltre cui si ritiene più probabile che le misure non seguano l'ipotesi nulla.

Dunque, a priori si fissa il livello di significatività o di confidenza, α , che definisce l'insieme dei valori che appartengono alla regione di rifiuto e di accettazione: α rappresenta la probabilità di commettere un errore di tipo I, ossia di respingere erroneamente l'ipotesi nulla quando è vera. Di solito, si utilizzano livelli di significatività comuni come 0,05 o 0,01.

Per stabilire se accettare o rifiutare l'ipotesi nulla, viene valutato il p-value: svolto un test statistico, che fornisce una misura di quanto i dati sperimentali siano in accordo con l'ipotesi nulla, il p-value è la probabilità di ottenere una statistica di test almeno pari a quella osservata, assumendo che l'ipotesi nulla sia vera. In sintesi, il p-value è un indicatore di quanto siano solidi i dati empirici per supportare o respingere un'ipotesi.

In base al confronto tra il p-value e il livello di significatività α , si prende una decisione: se il p-value è inferiore ad α , si rifiuta l'ipotesi nulla e si conclude che c'è evidenza statistica a favore dell'ipotesi alternativa; altrimenti, se il p-value è superiore ad α , non si rifiuta l'ipotesi nulla poiché i dati potrebbero essere coerenti con essa.

Il tempo di ritorno della precipitazione media areale massima annuale per il 2023 è stato stimato a partire dall'analisi della distribuzione dei valori sperimentali annuali ottenuti per i 32 anni di riferimento, dal 1991 al 2022.

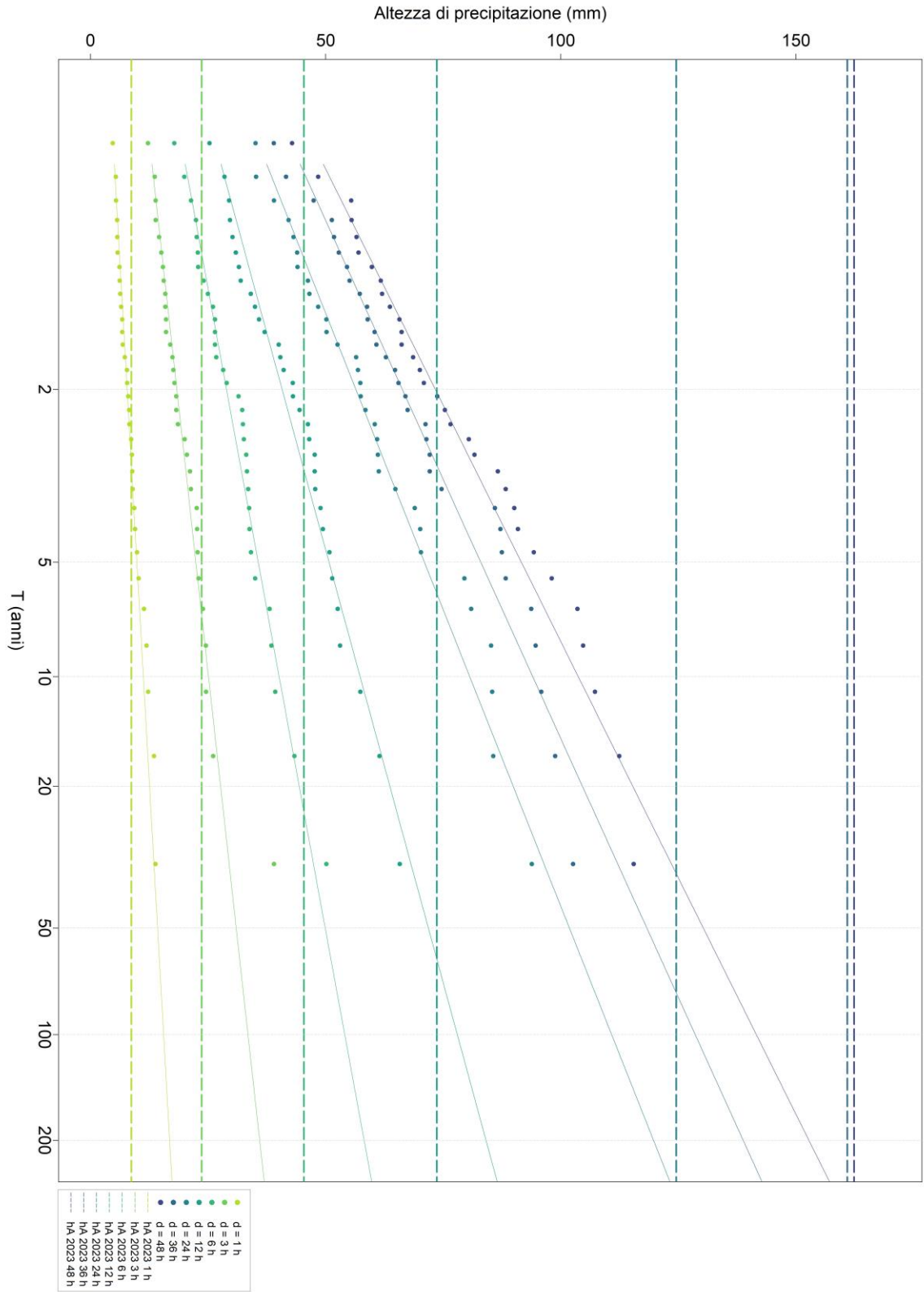
Seguendo la procedura e applicando i test statistici sopra descritti, automatizzati dalle funzionalità di RStudio, si è considerato il modello probabilistico di Gumbel, al fine di valutare, nel caso in cui i dati sperimentali seguano la distribuzione scelta, il tempo di ritorno delle altezze di precipitazione massima annuale media areale per varie durate relative al 2023, nel caso in cui risulti che gli eventi verificatisi ne abbiano determinato un carattere eccezionale.

A tal fine, sono stati elaborati i risultati sperimentali in termini di altezza media areale massima annuale, riferiti al periodo dal '91 al 2022: tramite la Weibull plotting position, ne è stata valutata la cumulata di non superamento empirica. I risultati ottenuti sono stati rappresentati su cartogramma probabilistico doppio esponenziale (o di Gumbel); si è osservato che i valori massimi della distribuzione sui 32 anni per le durate di 12, 24, 36 e 48 ore sono inferiori alla precipitazione media areale massima ricavata per il 2023. In riferimento a tali durate, si è proceduto con il confronto con la CDF teorica su carta di Gumbel. La carta probabilistica doppio esponenziale è un diagramma con scala delle ordinate deformata in modo tale che la funzione di probabilità cumulata di Gumbel sia rappresentata da una retta:

$$y = -\ln(-\ln(F_s(x_i)))$$

Graficamente, si è ottenuto per ciascuna macroarea (Figura 5.9 a e b):

Precipitazione media areale massima annuale in funzione della durata e del tempo di ritorno per la macro-area regionale Arpae A



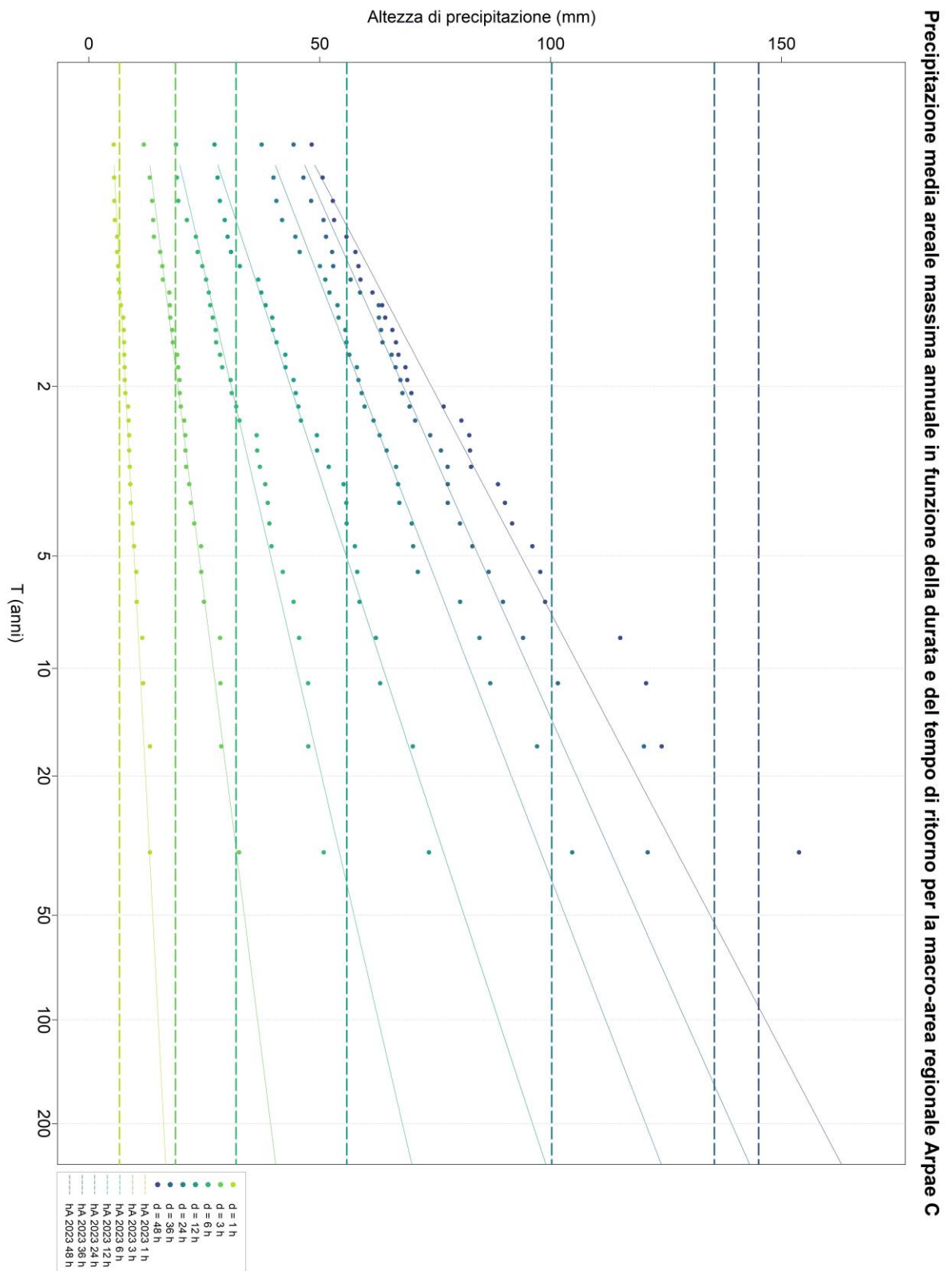


Figura 5.9 Precipitazione media areale massima annuale in funzione della durata e del tempo di ritorno per le macroaree A (a) e C (b), con rappresentazioni delle distribuzioni di Gumbel fittate sui dati sperimentali e indicazione dei valori sperimentali di $h_{A,max}(d)$ del 2023.

Sono stati eseguiti i PPCC test ed è stato calcolato il p-value con riferimento alle quattro curve teoriche fittate sui dati, in modo da valutarne la bontà di rappresentazione e la scelta del modello stesso.

	Macroarea A			
Durata (h)	12	24	36	48
PPCC	0,9844	0,9886	0,9808	0,9886
p-value	0,496	0,689	0,375	0,683

	Macroarea C
Durata (h)	36
PPCC	0,9896
p-value	0,7349

Si osserva quindi che le precipitazioni sperimentali medie areali si caratterizzano da tempi di ritorno notevoli, superiori ai 200 anni in corrispondenza di una durata di 12, 24, 36, 48 ore per la macroarea A, ma anche per una durata di 36 ore per la macroarea C. Questo, come illustrato nel paragrafo successivo, è dovuto alla natura stratiforme e alla persistenza dell'evento, sia nello spazio che nel tempo, con durata superiore ai tempi di corrivazione dei bacini (6-12 ore), che lo hanno reso eccezionale. La precipitazione media areale massima annuale, sperimentalmente è stata valutata come:

$$h_{A,max,j}(d = 1h) = \max_j \left(\sum_{i=1}^n P_i h_i(d = 1h) \right)$$

Ovvero è il valore massimo delle cumulate delle medie pesate delle precipitazioni puntuali, precipitazioni stesse che assumono valori notevoli per durate elevate, e quindi un risultato come quello relativo alla macroarea A si può spiegare in questo modo: per durate elevate, la precipitazione media sulla macroarea A, di 2500 km², raggiunge circa 160 mm (36 ore), con altezze puntuali massime che vanno da circa 80 a 240 mm (36 ore).

5.3 Boxplot

L'influenza degli eventi eccezionali che ha interessato l'Emilia-Romagna recentemente sulla valutazione del fattore di riduzione delle piogge è stata indagata tramite la rappresentazione della distribuzione dell'ARF annuale per fissata durata, costituita da 32 osservazioni, mediate boxplot; nel dettaglio, si è costruito un grafico (Figura 5.10) su cui sono state rappresentate le

scatole relative a ciascuna durata per le due macroaree e i punti relativi all'anno 2023, per ciascuna durata e macroarea considerate.

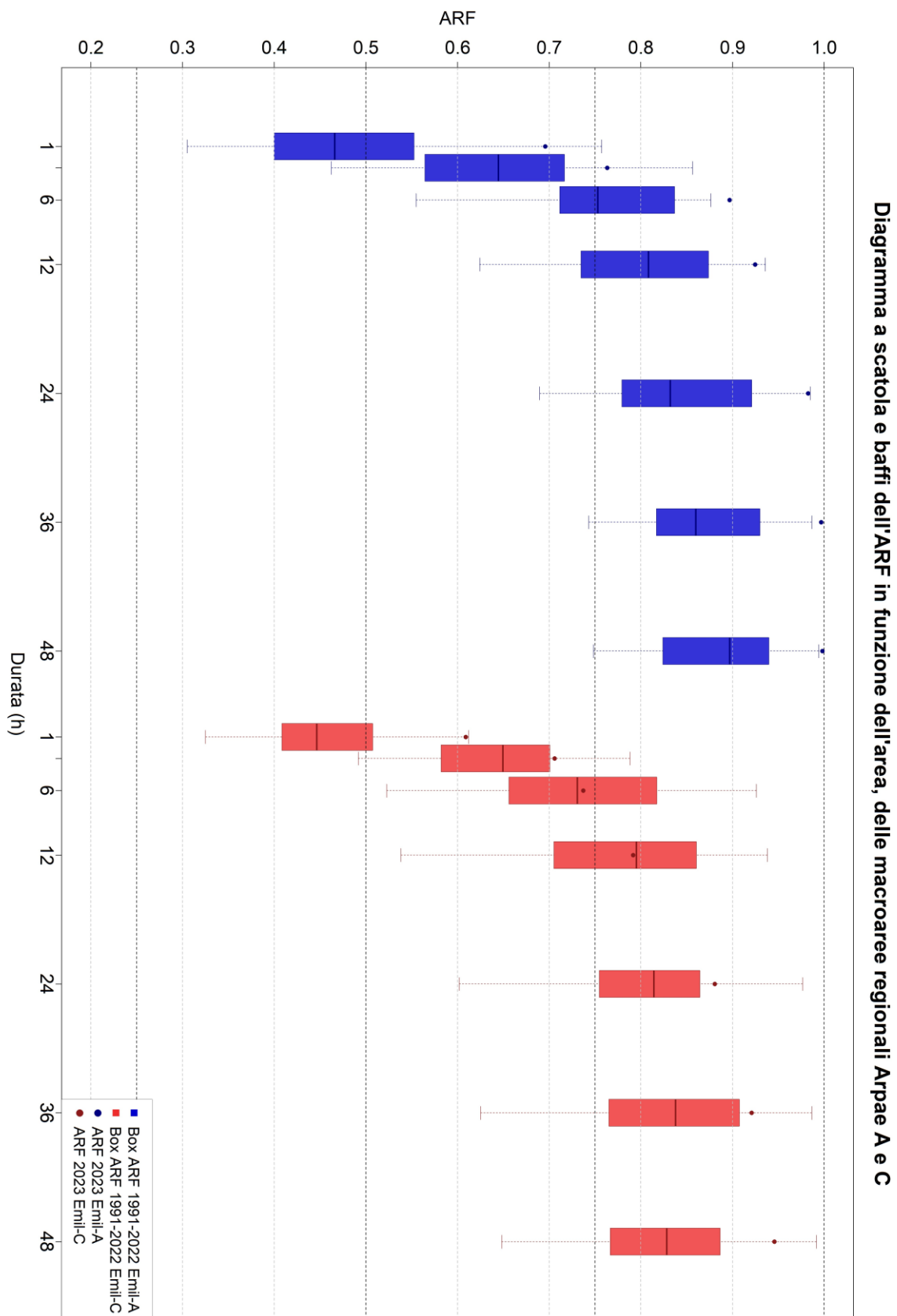


Figura 5.10 Boxplot della distribuzione su 32 anni (1991-2022) di valori sperimentali di ARF medio annuale, e ARF_{2023} , per ogni durata considerata e le due macroaree A e C.

Il caso di maggiore interesse, in concordanza con i risultati ottenuti in termini di altezza di precipitazione areale massima, è quello della macroarea A, per la quale si osserva che l'ARF₂₀₂₃ si colloca agli estremi, ovvero al di sopra del 75° percentile, delle distribuzioni dei valori relativi ai 32 anni precedenti per durate di 1, 3, 12 e 24 ore, andando a coincidere con il valore massimo della distribuzione nel caso di durate di 12 e 24 ore; mentre per durate di 6 e 36 e 48 ore, questo addirittura è un outlier, e cioè si discosta notevolmente dal comportamento tipico della distribuzione stessa.

Con riferimento invece alla macroarea C, di circa 3100 km², l'ARF₂₀₂₃ rientra nella distribuzione dei 32 anni precedenti, addirittura per brevi durate si posiziona sulla mediana, ed incrementa all'aumentare della durata, ma non ha un comportamento concorde con quanto si osserva per il primo caso.

Diversi studi di letteratura hanno indagato la dipendenza del coefficiente di ragguaglio dal tempo di ritorno della precipitazione puntuale, che si vuole spazializzare; molti di questi hanno portato alla conclusione che, per aree elevate, si rilevano ARF decrescenti con l'aumentare del periodo di ritorno, spiegando tale risultato come correlato al fatto che le precipitazioni areali diventano relativamente inferiori a causa dell'aumento dell'attività convettiva. Come ripetuto in precedenza, gli eventi convettivi tendono ad avere una minore estensione rispetto agli eventi di pioggia stratiformi, e quindi ci si aspettano decadimenti più forti dell'ARF con l'aumento dell'area.

Gli eventi alluvionali del maggio 2023 hanno però spiccata natura stratiforme, che spiega l'andamento al crescere della durata dell'ARF₂₀₂₃ rispetto agli anni precedenti per la macroarea A: in questo caso, le precipitazioni orografiche hanno determinato precipitazioni medie areali ad elevato tempo di ritorno, così come lo sono le precipitazioni cumulate puntuali massime alle quali vengono rapportate:

$$ARF(A, d) = \frac{h_{A,max,j}(d)}{h_{max,j,m}(d)}$$

Il risultato ricavato si può spiegare quindi tramite la genesi della precipitazione, il suo meccanismo, che si discosta dal comportamento standard dell'ARF annuale per un'area largamente estesa. La straordinaria persistenza dell'evento spaziale e temporale e la consecutività di due eventi estremi, che ha accentuato l'eccezionalità delle precipitazioni avvenute tra il 16 e il 18 maggio 2023, hanno determinato un valore del fattore di riduzione delle piogge all'area per la macroarea A, di circa 2500 km², che non rispetta il comportamento osservato nei 32 anni precedenti.

Quanto si osserva per la macroarea C si potrebbe invece imputare all'influenza del tempo di ritorno del massimo puntuale, come sopra descritto, che in questo caso è prevalente, data la precipitazione media areale generalmente inferiore al valore massimo storico dei 32 anni precedenti, e quindi anche alla maggiore estensione della macroarea stessa, ma soprattutto alla maggiore distanza dall'epicentro dell'evento. Alcuni studi (Korbinian et al.) individuano una possibile spiegazione ad ARF relativamente bassi per zone estese con predominanza generale di precipitazioni stratiformi, in effetti Stau particolarmente intensi, che possono portare a netti contrasti su piccola scala nei totali delle precipitazioni, come è tipico nei casi di precipitazioni orografiche.

D'altra parte, la dimensione del campione ridotta dovuta a meno di un anno di dati non consente di trarre conclusioni definitive.

In entrambi i casi, si tratta di porzioni del territorio fortemente colpite dagli eventi alluvionali di maggio, con precipitazioni cumulate puntuali e medie areali notevoli.

Conclusioni

Il presente lavoro si è posto l'obiettivo di calibrare alcune formulazioni di letteratura sui dati sperimentali di fattore di riduzione delle precipitazioni all'area valutati su bacini e sotto bacini dei principali corsi d'acqua appenninici emiliani.

La metodologia di stima dell'ARF proposta si basa sull'elaborazione dei massimi annuali e puntuali annuali, a partire da 32 anni di osservazioni a passo orario ottenute dal database meteorologico ERG5 di ARPAE. La procedura è stata tratta da analisi precedenti svolte sul solo bacino del Reno e basate sugli studi di Catalano et al. (1988). I dati valutati analiticamente per bacini e sottobacini del F. Reno, F. Taro, F. Trebbia, F. Nure, per aree che variano da 25 ai 3500 km², sono coerenti con studi precedenti che hanno riscontrato che l'ARF diminuisce con l'aumento dell'area e con la diminuzione della scala temporale di aggregazione delle precipitazioni, comunemente indicata anche come durata.

I risultati ottenuti nell'elaborazione del dataset utilizzando questa tecnica sono stati confrontati con alcune delle più importanti equazioni empiriche di letteratura per il coefficiente di ragguaglio all'area. Si è quindi ottenuta una formulazione che descrivesse l'andamento dell'ARF per i bacini appenninici considerati attraverso la calibrazione sui dati sperimentali, con riferimento a tutte le durate considerate nello studio (1, 3, 6, 12, 24, 36, 48 ore). Infine, considerando le metriche di prestazione selezionate, si è verificato che la formulazione che meglio interpreta i dati sperimentali è quella proposta da Moisello e Papiri (1986) sebbene anche le curve ricalibrate dell'U.S. Weather Bureau (1972) e di Columbo (1960) forniscano risultati più che soddisfacenti. D'altra parte, è stato riscontrato che queste formulazioni non rappresentano efficacemente l'ARF stimato per brevi periodi e aree elevate, per le quali si trovano differenze non trascurabili.

Il coefficiente di ragguaglio risulta, in ogni caso, funzione dell'area, della durata ed anche del dominio spazio temporale di riferimento.

Nelle formulazioni di letteratura spesso viene considerata trascurabile l'influenza sul fattore di riduzione delle precipitazioni all'area dell'entità dell'evento pluviometrico, o meglio del tempo di ritorno. Detto aspetto non è stato verificato nel presente lavoro, ma sarà oggetto di approfondimenti successivi sulla base dei valori empirici dell'ARF ricavati in questo studio. Le indagini sviluppate hanno anche permesso di approfondire persistenza, uniformità spaziale e frequenza dei valori massimi delle precipitazioni medie areali per fissata durata caratteristiche dei recenti eventi alluvionali che hanno interessato la Romagna nel maggio del

2023. In particolare, la procedura sperimentale adottata è stata impiegata per la valutazione del coefficiente di ragguaglio in funzione della durata e dell'area con riferimento a due delle macroaree regionali ARPAE, A e C, ricoprenti la porzione occidentale dell'appennino emiliano; e a 32 anni di osservazioni, dal '91 al 2022. I recenti eventi alluvionali estremi che hanno colpito l'Emilia-Romagna sono stati presi in esame tramite il calcolo sulle macroaree dell'ARF_Y con riferimento ai primi mesi del 2023.

Dall'analisi statistica delle precipitazioni medie areali massime annuali e dal confronto, tramite opportuni diagrammi diagnostici, tra il risultato ottenuto per il 2023 e quanto invece ricavato per i 32 anni precedenti, l'effetto della natura delle precipitazioni e in particolare della loro uniformità spaziale su vaste aree, e la frequenza attesa dei massimi di precipitazione areale a fissata durata. In un caso si è osservato che la notevole persistenza nello spazio e nel tempo degli eventi dal 1° al 3 e dal 16 al 18 maggio 2023 hanno determinato una precipitazione media areale particolarmente significativa e sicuramente eccezionale in relazione ai precedenti 32 anni di osservazione, per durate elevate (12, 24, 36, 48 ore), con stime del periodo di ritorno crescenti con la durata, fino a valori ben superiori a 200 anni. La natura stratiforme della precipitazione, caratterizzata da intensità orarie non elevate ma da precipitazioni aggregate per lunghe durate eccezionali, al di sopra dei record storici mensili e annuali, e da una distribuzione su larga scala e scarsa variabilità spaziale, ha determinato una precipitazione media areale notevole. Questo risultato è ben evidente in riferimento alla macroarea A, più orientale.

Con riferimento alla più occidentale macroarea C, l'analisi ha portato a risultati sicuramente significativi, anche se con riferimento a questa macroarea sia i valori massimi di precipitazione areale che i valori sperimentali dell'ARF ottenuti per il maggio del 2023 non assumono carattere di eccezionalità rispetto al precedente trentennio di osservazioni. Si è supposto che l'estensione superiore della macroarea C, circa 3500 km², e la sua maggiore distanza dall'epicentro dell'evento, abbiano in parte determinato un coefficiente di ragguaglio che rientra nella distribuzione e che per durate intermedie si attesta sul valore mediano. Una possibile spiegazione è fornita dalla complessità delle precipitazioni orografiche, cui ha contribuito un effetto Stau, e dalla loro variabilità su diverse scale spaziali. Il risultato, inoltre, è concorde con quanto rilevato da studi precedenti in merito all'influenza del tempo di ritorno, secondo i quali il fattore di riduzione areale decresce per precipitazioni ad elevato T ed aree elevate.

È stata quindi evidenziata la forte dipendenza del coefficiente di ragguglio areale delle piogge dalla durata dell'evento, dall'area considerata, in accordo con gli studi classici di letteratura, ma anche dall'entità della precipitazione e dal suo meccanismo, come riportato da analisi più recenti.

Appendice

RStudio script per la valutazione dell'ARF annuale

In questa sezione vengono raccolti gli script realizzati per l'esecuzione dei vari passaggi della procedura di valutazione dell'ARF annuale adottata nel lavoro illustrato:

- Database ERG5;
- Dataset di riferimento;
- Elaborazione dei file SHP e selezione delle celle del grigliato ERG5 che ricadono nell'area di interesse;
- Filtro per periodo di riferimento per il calcolo;
- Aggregazione delle precipitazioni a passo orario alle durate considerate (1, 3, 6, 12, 24, 36, 48 h);
- Estrazione dei massimi annuali delle singole celle e calcolo del valore medio massimo annuale di precipitazione per ogni durata;
- Media areale sull'area considerata, calcolo della media areale cumulata secondo ciascuna durata ed estrazione del valore massimo annuale di precipitazione media areale;
- Valutazione del fattore di riduzione delle piogge all'area, ARF.

In cui sono riportati anche i passaggi aggiuntivi per l'elaborazione dell'ARF con riferimento dell'evento alluvionale verificatosi nel maggio 2023 in Emilia Romagna. In tale ambito, è presente anche lo script per la rappresentazione dei massimi annuali di precipitazione media areale e la valutazione del percentile corrispondente ai valori di ARF stimati per il 2023.

```

# ARCHIVIO BASE - Database ERG5
# Installazione e caricamento pacchetti necessari
install.packages("lubridate")
install.packages("tidyverse")
library(lubridate)
library(tidyverse)
# Definizione del percorso della cartella contenente il
dataset suddiviso
# in due sottoperiodi
file_path <-
list.files("C:/Users/giuditta.smerilli/Desktop/ARF/Dataset",
full.names = TRUE)
# Definizione oggetti data_iniziale e data_finale per la
generazione della
# dataframe di NA per controllare la presenza di NA nel
dataset
data_iniziale <- ymd_hms("1991-01-01 01:00:00", tz = "UTC")
data_finale <- ymd_hms("2023-05-22 23:00:00", tz = "UTC")
# Inizializzazione dataframe di NA con un numero di righe pari
al numero di ore
# e un numero di colonne pari al numero di celle e una colonna
per la data
# Viene generata una sequenza di ore utilizzando la funzione
seq e i limiti
# di data specificati nel vettore data_iniziale e data_finale
ore <- seq(from = data_iniziale, to = data_finale, by =
"hour")
numero_ore <- length(ore)
# Vengono elencati i file nella cartella corrente utilizzando
# la funzione list.files con il pattern ".csv"
for (i in 1 : length(file_path))
{
  files <- list.files(file_path[i], pattern = ".csv",
full.names = TRUE)
  assign(paste0("files_", i), files)
}
numero_file <- length(files_1)
df_global <- data.frame(matrix(rep(NA,
(numero_ore)*(numero_file + 1)), nrow = numero_ore))
df_global[, 1] <- ore
for (j in 1 : numero_file)
{
  data <- data.frame()
  for (i in 1 : length(file_path))
  {
    nome_files <- paste0("files_", i)
    files <- get(nome_files)
    # Per ogni file, si usa la function read.csv per leggerne
il contenuto nel dataframe dati
    dati <- read.csv(files[j], sep = ",", dec = ".", header =
TRUE)

```

```

# Viene applicato il filtro al dataframe data per
selezionare solo le righe in cui
# la colonna "VARIABLE" ha il valore "PREC" (nel dataframe
dal 2001 al 2023, sono riportati
# oltre al valore dell'altezza di pioggia, anche
temperatura, radiazione, vento)
dati <- dati %>%
  filter(VARIABLE == "PREC")
data <- rbind(data, dati)
}
# Le colonne del dataframe df successive alla prima colonna
vengono assegnate con i valori
# della colonna 4 del dataframe data (VARIABLE)
df_global[, j + 1] <- data[, 4]
}
# Vengono assegnati i nomi delle colonne del dataframe df
codice <- gsub(".*/(\\d{5})\\.csv", "\\1", files_1)
nome <- paste0("PREC_", codice)
colnames(df_global) <- c("DATA", nome)
# Viene eseguito un controllo sulla presenza di valori
mancanti (NA) nel dataframe df
if (any(is.na(df_global))) {
  print("Sono presenti valori mancanti (NA) nel dataframe.")
} else {
  print("Non ci sono valori mancanti (NA) nel dataframe.")
}
# Salvataggio del df creato in formato R Data
save(df_global, file = "df_1991_2023.rda")
# Apertura del df contenente dati di precipitazione dal 1991-
01-01 01:00:00 al 2023-05-22 23:00:00
path_df <- "C:/Users/giuditta.smerilli/OneDrive - Alma Mater
Studiorum Università di Bologna/ARF/R/df_1991_2023.rda"
load(path_df)

```

```

# FILTRO per periodo
# Installazione e caricamento pacchetti necessari
install.packages("lubridate")
install.packages("tidyverse")
library(lubridate)
library(tidyverse)
# Selezione delle serie orarie relative al bacino e agli anni
di interesse
# Definizione vettore anni considerati
Anni_ERG5 <- 1991:2022
# Filtro per anno
df_global <- df_global %>% mutate(ANNO = year(DATA)) %>%
relocate(ANNO, .after = DATA) %>%
  filter(ANNO %in% Anni_ERG5)
# Floods
Anno_floods <- 2023
# Filtro per anno, caso floods
df_global <- df_global %>% mutate(ANNO = year(DATA)) %>%
relocate(ANNO, .after = DATA) %>%
  filter(ANNO %in% Anno_floods)

```



```

# FILTRO per area, calcolo dei PESI
# Installazione e caricamento pacchetti necessari
install.packages("terra")
install.packages("lubridate")
install.packages("tidyverse")
library(terra)

# Elaborazione shapefile per il calcolo delle aree di
# pertinenza delle celle nel bacino considerato
# Definizione di oggetti stringa con i path degli shapefile
# grigliato e area
path_area <- "C:/Users/giuditta.smerilli/OneDrive - Alma Mater
Studiorum Università
di Bologna/ARF/QGIS/Bacini_all/all_288_bacini.shp"
path_grigliato <- "C:/Users/giuditta.smerilli/OneDrive - Alma
Mater Studiorum Università
di Bologna/ARF/QGIS/GRIGLIATO/Grigliato_ERG5.shp"
# Apertura degli shapefile vettoriali area e grigliato
Area <- vect(path_area)
Grigliato <- vect(path_grigliato)
# Si rinomina il primo campo dello spatvector Area come "NOME"
names(Area)[3] <- "NOME"
# Selezione del bacino di interesse e dei suoi sottobacini
bac_glob <- Area[Area$NOME == "shp_B019", ] # Reno shp_B019,
Trebbia shp_B087, Nure shp_B076
Area <- crop(Area, bac_glob)
# Caso del Taro :
# pol_Taro <- c("shp_A176", "shp_B067", "shp_B041",
"shp_B092")
# bac_glob <- Area[Area$NOME == "shp_A176", ]
# for (w in 2 : length(pol_Taro)) bac_glob <- union(bac_glob,
Area[Area$NOME == pol_Taro[w], ])
# Suddivisione del bacino in poligoni separati sulla base del
# nome
lista_SHP <- split(Area, "NOME")
id_all <- NULL
for (j in 1:length(lista_SHP))
{
  Area_j <- lista_SHP[[j]]
  Area_j_kmq <- expanse(Area_j, unit="km")
  # Selezione dei bacini in funzione della loro estensione, in
  # modo che si abbiano bacini con area
  # superiore all'area delle celle del Grigliato (25 km2)
  if (Area_j_kmq >= 25)
  {
    id_single <- Area_j$NOME
    id_all <- c(id_all, id_single)
    # Assegnazione dello stesso sistema di coordinate di
    # riferimento CRS prima di eseguire l'intersezione
    Grigliato <- project(Grigliato, crs(Area_j))
    # Ritaglio del grigliato rispetto al bacino utilizzando la

```

```

funzione crop
  Area_j_cut <- crop(Grigliato, Area_j)
  # Calcolo dell'area di ogni poligono nell'oggetto
Bacino_cut utilizzando la funzione expanse()
  # e assegnando il risultato al campo "Area_kmq"
  Area_j_cut$Area_kmq <- expanse(Area_j_cut, unit="km")
  # Calcolo dei pesi
  Area_j_cut$Peso <- Area_j_cut$Area_kmq /
sum(Area_j_cut$Area_kmq)
  # Creazione di un dataframe contenente la tabella degli
attributi finale con le informazioni
  # delle celle del grigliato che ricadono nel bacino
  Tab_attributi <- data.frame(Area_j_cut)
  # Se ci sono più poligoni nello shapefile, si assegna il
dataframe Tab_attributi a un oggetto
  # con un nome univoco basato sul nome del poligono
  if (length(lista_SHP) > 1)
    {
      name <- Area_j$NOME
      assign(paste0("Tab_attributi_", name), Tab_attributi)
    }
}
}
id_all <- sort(id_all)
# Selezione della tabella degli attributi del poligono di
interesse, nel caso in cui lo shapefile in input sia
# multipoligonale e si voglia prendere in esame solo uno dei
poligoni in esso presenti o selezionarne più di uno
if (length(lista_SHP) > 1)
{
  id_selez <- id_all[20]
  nome_tab <- paste0("Tab_attributi_", id_selez)
  Tab_attributi <- get(nome_tab)
}
# Definizione del vettore contenente i codici delle celle che
ricadono nel bacino e quindi del vettore contenente
# i nomi delle colonne di interesse
Cella <- Tab_attributi$Codice
library(lubridate)
library(tidyverse)
nomi <- sprintf("PREC_%05d", Cella)
# Filtro per bacino
df <- df_global %>% select(DATA, ANNO, all_of(nomi))

```

```

# AGGREGAZIONE PRECIPITAZIONE
# Creazione del vettore durata contenente le durate per le
quali calcolare le cumulate
durata <- c(1, 3, 6, 12, 24, 36, 48)
df_cumulata_1h <- df
# Inizializzazione del dataframe df_cumulata
col_df_data <- df[1 : nrow(df), 1]
col_df <- df[, 3 : ncol(df)]
# Ciclo for che itera su ogni elemento di durata
for (iddur in 2 : length(durata)) #par forse meglio
{
  df_cumulata <- data.frame(matrix(rep(NA, (nrow(df) -
durata[iddur] + 1) * (ncol(df))),
                                nrow = nrow(df) -
durata[iddur] + 1))
  col_df_cum <- NULL
  index_dur <- NULL
  for (idr in 1 : nrow(df_cumulata)) # no par
  {
    finestra <- col_df[idr : (idr + durata[iddur] - 1),]
    col_df_cum_r <- cumsum(finestra)
    col_df_cum_r <- col_df_cum_r[nrow(finestra),]
    col_df_cum <- rbind(col_df_cum, col_df_cum_r)
    index_r <- idr + durata[iddur] - 1
    index_dur <- c(index_dur, index_r)
    print(paste0("Durata = ", durata[iddur],"; riga = ",
idr))
  }
  col_df_dur <- col_df_data[index_dur]
  df_cumulata[, 1] <- col_df_dur
  ANNO <- year(col_df_dur)
  df_cumulata[, 2] <- ANNO
  df_cumulata[, 3 : ncol(df_cumulata)] <- col_df_cum[, 1 :
ncol(col_df_cum)]
  colnames(df_cumulata) <- colnames(df)
  nome_df <- paste0("df_cumulata_", durata[iddur], "h")
  assign(nome_df, df_cumulata)
}

```

```

# ESTRAZIONE dei massimi annuali
install.packages(tools)
library(tools)
# Selezione della cartella per il salvataggio dei CSV
contenenti i massimi annuali precipitazione
# cumulata di ciascuna cella, i CSV con le date in cui si sono
relizzati questi valori massimi,
# i CSV con la media pesata dei valori massimi annuali di
precipitazione relativa alle singole celle
cartella_max_annual_all <-
"C:/Users/giuditta.smerilli/OneDrive - Alma Mater Studiorum
Università
di Bologna/ARF/R/Massimi_annuali_celle_"
cartella_max_annual_wd_all <-
"C:/Users/giuditta.smerilli/OneDrive - Alma Mater Studiorum
Università
di Bologna/ARF/R/Massimi_annuali_celle_wd_"
cartella_max_annual_mean_all <-
"C:/Users/giuditta.smerilli/OneDrive - Alma Mater Studiorum
Università
di Bologna/ARF/R/Massimi_annuali_celle_mean_"
name_area <- c("Macro_A_C", "Macro_A_C_2023", "Nure", "Reno",
"Taro", "Trebbia")
name_area_sel <- name_area[4]
cartella_max_annual <- paste0(cartella_max_annual_all,
name_area_sel, "/")
cartella_max_annual_wd <- paste0(cartella_max_annual_wd_all,
name_area_sel, "/")
cartella_max_annual_mean <-
paste0(cartella_max_annual_mean_all, name_area_sel, "/")
for (i in durata)
{
  nome_df <- paste0("df_cumulata_", i, "h")
  df_cum <- get(nome_df)
  # Estrazione dei massimi annuali delle singole celle per
ogni durata
  max_annual <- df_cum %>%
  group_by(ANNO) %>%
  summarize_at(vars(starts_with("PREC_")), ~ max(., na.rm =
TRUE)) %>% ungroup()
  nome <- paste0("max_annual_", i, "h")
  # Salvataggio dei CSV
  if (length(lista_SHP) > 1) {
    write.csv(x = max_annual, file =
paste0(cartella_max_annual, nome, "_",
file_path_sans_ext(basename(path_area)),
"_", id_selez,
".csv"), sep = ",", dec = ".")
  } else {
    write.csv(x = max_annual, file =

```

```

paste0(cartella_max_annual, nome, "_",
file_path_sans_ext(basename(path_area))),
      ".csv", sep = ",", dec = ".")
  }
  assign(nome, max_annual)
  df_max <- get(nome)
  # Estrazione delle date in cui si sono verificati i massimi
  # Immagazzinamento in un dataframe in cui sono riportati i
massimi e le corrispondenti date
  max_annual_withdate <- data.frame(ANNO = df_max$ANNO)
  for (j in 1:(ncol(df_max) - 1))
    {
      col_name_max <- colnames(df_max)[j+1]
      col_name_date <- paste0("DATA_", substr(col_name_max, 6,
10))
      idx_max <- match(df_max[[col_name_max]],
df_cum[[col_name_max]])
      DATA <- df_cum$DATA[idx_max]
      MAX <- df_max[[col_name_max]]
      if (j == 1) {
        max_annual_withdate <- data.frame(max_annual_withdate,
DATA, MAX)

colnames(max_annual_withdate)[2:length(max_annual_withdate)]
<- c(col_name_date, col_name_max)
      } else {
        new_col <- data.frame(DATA, MAX)
        colnames(new_col) <- c(col_name_date, col_name_max)
        max_annual_withdate <- cbind(max_annual_withdate,
new_col)
      }
    }
  nome_mwd <- paste0("max_annual_withdate_", i, "h")
  # Salvataggio in CSV
  if (length(lista_SHP) > 1) {
    write.csv(x = max_annual_withdate, file =
paste0(cartella_max_annual_wd, nome_mwd, "_",
file_path_sans_ext(basename(path_area)), "_",
      id_selez,
".csv"), sep = ",", dec = ".")
  } else {
    write.csv(x = max_annual_withdate, file =
paste0(cartella_max_annual_wd, nome_mwd, "_",
file_path_sans_ext(basename(path_area))),
      ".csv", sep = ",", dec = ".")
  }
  assign(nome_mwd, max_annual_withdate)
}

```

```

# Calcolo della media pesata dei valori massimi annuali per le
singole celle
for (i in durata)
{
  nome_df <- paste0("max_annual_", i, "h" )
  df_max <- get(nome_df)
  df_max_mean <- df_max %>%
    rowwise() %>%
    mutate(PREC_MAXP_MEAN =
weighted.mean(c_across(starts_with("PREC_")), w =
Tab_attributi$Peso) %>%
  select(ANNO, PREC_MAXP_MEAN)
  nome_df_maxmean <- paste0("mean_max_annual_", i, "h")
  # Salvataggio in CSV
  if (length(lista_SHP) > 1)
  {
    write.csv(x = df_max_mean, file =
paste0(cartella_max_annual_mean, nome_df_maxmean, "_",
file_path_sans_ext(basename(path_area)), "_",
id_selez,
".csv"), sep = ",", dec = ".")
  } else {
    write.csv(x = df_max_mean, file =
paste0(cartella_max_annual_mean, nome_df_maxmean, "_",
file_path_sans_ext(basename(path_area))),
".csv", sep = ",", dec = ".")
  }
  assign(nome_df_maxmean, df_max_mean)
}

```

```

# MEDIA AREALE
# Calcolo media areale a partire dai dati a passo orario
media_areale <- df %>%
  rowwise() %>%
  mutate(PREC_MA =
weighted.mean(c_across(starts_with("PREC_")), w =
Tab_attributi$Peso)) %>%
  select(DATA, ANNO, PREC_MA)
# Cumulata secondo le durate definite nel vettore durata
col_ma_data <- media_areale[, 1]
col_ma <- media_areale[, 3]
for (i in 1:length(durata))
{
  media_areale_cum <- data.frame(matrix(rep(NA,
(nrow(media_areale) - durata[i] + 1) * (ncol(media_areale))),
nrow =
nrow(media_areale) - durata[i] + 1))
  col_ma_cum <- NULL
  index_dur <- NULL
  for (r in 1 : nrow(media_areale_cum))
  {
    finestra <- col_ma[r : (r + durata[i] - 1), ]
    col_ma_cum_r <- sum(finestra)
    col_ma_cum <- c(col_ma_cum, col_ma_cum_r)
    index_r <- r + durata[i] - 1
    index_dur <- c(index_dur, index_r)
  }
  media_areale_cum[, 3] <- col_ma_cum
  col_ma_dur <- col_ma_data[index_dur,]
  media_areale_cum[, 1] <- col_ma_dur[, 1]
  ANNO <- media_areale[index_dur, 2]
  media_areale_cum[, 2] <- ANNO[, 1]
  colnames(media_areale_cum) <- colnames(media_areale)
  nome <- paste0("media_areale_cum_", durata[i], "h")
  assign(nome, media_areale_cum)
}
# Definizione cartella di destinazione
name_area <- c("Macro_A_C", "Macro_A_C_2023", "Nure", "Reno",
"Taro", "Trebbsia")
name_area_sel <- name_area[4]
cartella_max_medareali_annual_all <-
"C:/Users/giuditta.smerilli/OneDrive - Alma Mater
Studiorum Università di
Bologna/ARF/R/Massimi_annuali_medie_areali_"
cartella_max_medareali_annual <-
paste0(cartella_max_medareali_annual_all, name_area_sel, "/")
# Estrazione massimi annuali areali
for (i in durata)
{
  nome_df <- paste0("media_areale_cum_", i, "h")
  df_ma_cum <- get(nome_df)
}

```

```

max_annual_ma <- df_ma_cum %>%
  group_by(ANNO) %>% summarize(PREC_MAXMA = max(PREC_MA,
na.rm = TRUE)) %>%
  ungroup() %>% select(ANNO, PREC_MAXMA)
nome <- paste0("max_annual_ma_", i, "h")
# Salvataggio in CSV
if (length(lista_SHP) > 1) {
  write.csv(x = max_annual_ma, file =
paste0(cartella_max_medareali_annual,
file_path_sans_ext(basename(path_area)),
nome, "_",
"_", id_selez,
".csv"), sep = ",", dec = ".")
} else {
  write.csv(x = max_annual_ma, file =
paste0(cartella_max_medareali_annual,
nome, "_",
file_path_sans_ext(basename(path_area))),
".csv", sep = ",", dec = ".")
}
assign(nome, max_annual_ma)
}

```



```

# Calcolo ARF
# Definizione cartella destinazione CSV
cartella_ARF_annual_all <-
"C:/Users/giuditta.smerilli/OneDrive - Alma
Mater Studiorum Università di Bologna/ARF/R/ARF_annuale_"
name_area <- c("Macro_A_C", "Macro_A_C_2023", "Nure", "Reno",
"Taro", "Trebbia")
name_area_sel <- name_area[4]
cartella_ARF_annual <- paste0(cartella_ARF_annual_all,
name_area_sel, "/")
# Calcolo dell'ARF annuale per ogni durata considerata
for (i in durata)
{
  nome_max <- paste0("mean_max_annual_", i, "h")
  max_p <- get(nome_max)
  nome_ma <- paste0("max_annual_ma_", i, "h")
  ma_max <- get(nome_ma)
  ARF <- data.frame(ANNO = max_p$ANNO, ARF = ma_max$PREC_MAXMA
/ max_p$PREC_MAXP_MEAN)
  nome_arf <- paste0("ARF_", i, "h")
  # Salvataggio in CSV
  if (length(lista_SHP) > 1)
  {
    write.csv(x = ARF, file = paste0(cartella_ARF_annual,
nome_arf, "_",
file_path_sans_ext(basename(path_area)), "_", id_selez,
".csv"), sep = ",", dec =
".")
  } else {
    write.csv(x = ARF, file = paste0(cartella_ARF_annual,
nome_arf, "_",
file_path_sans_ext(basename(path_area))),
".csv", sep = ",", dec = ".")
  }
  assign(nome_arf, ARF)
}

```

RStudio script delle functions delle formule di letteratura calibrate

Sono raccolti anche gli script impiegati in sede di calibrazione delle formule di letteratura considerate nell'analisi dei risultati ottenuti:

- Function corrispondente alla formulazione dell'USWB, preceduta dall'inizializzazione delle variabili interne, comune alle altre funzioni; e funzioni di appoggio per le metriche impiegate nella calibrazione.
- Function corrispondente alla formulazione di Moisello e Papiri;
- Function corrispondente alla formulazione di Columbo.

```

# funzione USWB
library("terra")
durata <- c(1, 3, 6, 12, 24, 36, 48)
folder_ARF_base <- "C:/Users/giuditta.smerilli/OneDrive - Alma
Mater Studiorum Università di Bologna/ARF/R/ARF_annuale_"
path_area <- "C:/Users/giuditta.smerilli/OneDrive - Alma Mater
Studiorum Università di
Bologna/ARF/QGIS/Bacini_all/all_288_bacini.shp"
Area <- vect(path_area)
names(Area)[3] <- "NOME"
name_area <- c("Nure", "Reno", "Taro", "Trebbia")
name_pol <- c("shp_B076", "shp_B019", "shp_A176", "shp_B087")
pol_Taro <- c("shp_A176", "shp_B067", "shp_B041", "shp_B092")
A <- NULL
for (j in 1 : length(name_area))
{
  id_all <- NULL
  Area_kmq <- NULL
  bac_glob <- Area[Area$NOME == name_pol[j], ]
  if (name_area[j] == "Taro") {
    for (u in 2 : length(pol_Taro)) bac_glob <-
union(bac_glob, Area[Area$NOME == pol_Taro[u], ])
  } else { bac_glob <- Area[Area$NOME == name_pol[j], ] }
  Area_crop <- crop(Area, bac_glob)
  lista_SHP <- split(Area_crop, "NOME")
  for (s in 1:length(lista_SHP)) {
    Area_j <- lista_SHP[[s]]
    Area_j_kmq <- expanse(Area_j, unit="km")
    if (Area_j_kmq >= 25) {
      id_single <- Area_j$NOME
      id_all <- c(id_all, id_single)
      Area_kmq <- c(Area_kmq, Area_j_kmq)
    }
  }
  index <- order(id_all)
  id_all <- id_all[index]
  A <- c(A, Area_kmq[index])
}
ARF_oss <- data.frame(matrix(rep(NA, length(A) *
(length(durata) + 1)), nrow = length(A)))
ARF_oss[, 1] <- A
for (i in 1 : length(durata))
{
  col_dur <- NULL
  for (n in 1 : length(name_area))
  {
    ARF_mean_dur_i <- NULL
    file_path_dur_i <- list.files(path =
paste0(folder_ARF_base, name_area[n]), pattern = paste0("_",
durata[i], "h_"), full.names = TRUE)
    for (z in 1 : length(file_path_dur_i)) {
      ARF_dur_area_p <- read.csv(file_path_dur_i[z], sep =

```

```

",", dec = ".", header = TRUE)
  ARF_mean_dur_i[z] <- mean(ARF_dur_area_p[, 3])
}
col_dur <- c(col_dur, ARF_mean_dur_i)
}
ARF_oss[, i + 1] <- col_dur
}
ARF <- data.frame(matrix(rep(NA, length(A) * (length(durata) +
1)), nrow = length(A)))
ARF[, 1] <- A
ARF_oss_vec <- NULL
ARF_vec <- NULL
for (k in 2 : ncol(ARF_oss))
{
  ARF_oss_k <- ARF_oss[, k]
  ARF_oss_vec <- c(ARF_oss_vec, ARF_oss_k)
}
ARF_US <- function(parametri) {
  c1 <- parametri[1]
  c2 <- parametri[2]
  c3 <- parametri[3]

  for (p in 2 : ncol(ARF))
  {
    ARF[, p] <- 1 - (1 - exp(c1 * A)) * exp(c2 * ((durata[p -
1]) ^ c3))
    ARF_k <- ARF[, p]
    ARF_vec <- c(ARF_vec, ARF_k)
  }
  SSQ <- sum((ARF_oss_vec - ARF_vec)^ 2)
  N <- length(ARF_oss_vec)
  NSE <- (1 - SSQ / (var(ARF_oss_vec) * (N - 1)))
  p <- length(parametri)
  NSEadj <- 1 - (N - 1) / (N - p) * (1 - NSE)
  R2 <- cor(ARF_oss_vec, ARF_vec)^ 2
  R2adj <- 1 - (N - 1) / (N - p) * (1 - R2)
  ratio <- NULL
  for (t in 2 : ncol(ARF)) {
    ARF_mod <- ARF[, t]
    ARF_sper <- ARF_oss[, t]
    ratio_t <- sum((ARF_mod - ARF_sper) / ARF_sper)
    ratio <- c(ratio, ratio_t)
  }
  BIAS_rel <- (1 / length(durata)) * (1 / length(A)) *
sum(ratio)
  OUTPUT <- list(SSQ, NSE, NSEadj, R2, R2adj, BIAS_rel)
  return(OUTPUT)
}
# Funzioni di appoggio
NSEadj_US <- function(parametri) {
  A <- ARF_US(parametri)

```

```

    return(A[[3]])
}
RQadj_US <- function(parametri) {
  A <- ARF_US(parametri)
  return(A[[5]])
}

# funzione Columbo
ARF_COL <- function(parametri) {
  c1 <- parametri[1]
  c2 <- parametri[2]
  c3 <- parametri[3]
  c4 <- parametri[4]
  for (p in 2 : ncol(ARF))
  {
    ARF[, p] <- (1 - c1 * (A ^ c2)) * (durata[p - 1] ^ (c3 *
(A ^ c4)))
    ARF_k <- ARF[, p]
    ARF_vec <- c(ARF_vec, ARF_k)
  }
  SSQ <- sum((ARF_oss_vec - ARF_vec)^ 2)
  N <- length(ARF_oss_vec)
  NSE <- (1 - SSQ / (var(ARF_oss_vec) * (N - 1)))
  p <- length(parametri)
  NSEadj <- 1 - (N - 1) / (N - p) * (1 - NSE)
  R2 <- cor(ARF_oss_vec, ARF_vec)^ 2
  R2adj <- 1 - (N - 1) / (N - p) * (1 - R2)
  ratio <- NULL
  for (t in 2 : ncol(ARF)) {
    ARF_mod <- ARF[, t]
    ARF_sper <- ARF_oss[, t]
    ratio_t <- sum((ARF_mod - ARF_sper) / ARF_sper)
    ratio <- c(ratio, ratio_t)
  }
  BIAS_rel <- (1 / length(durata)) * (1 / length(A)) *
sum(ratio)
  OUTPUT <- list(SSQ, NSE, NSEadj, R2, R2adj, BIAS_rel)
  return(OUTPUT)
}

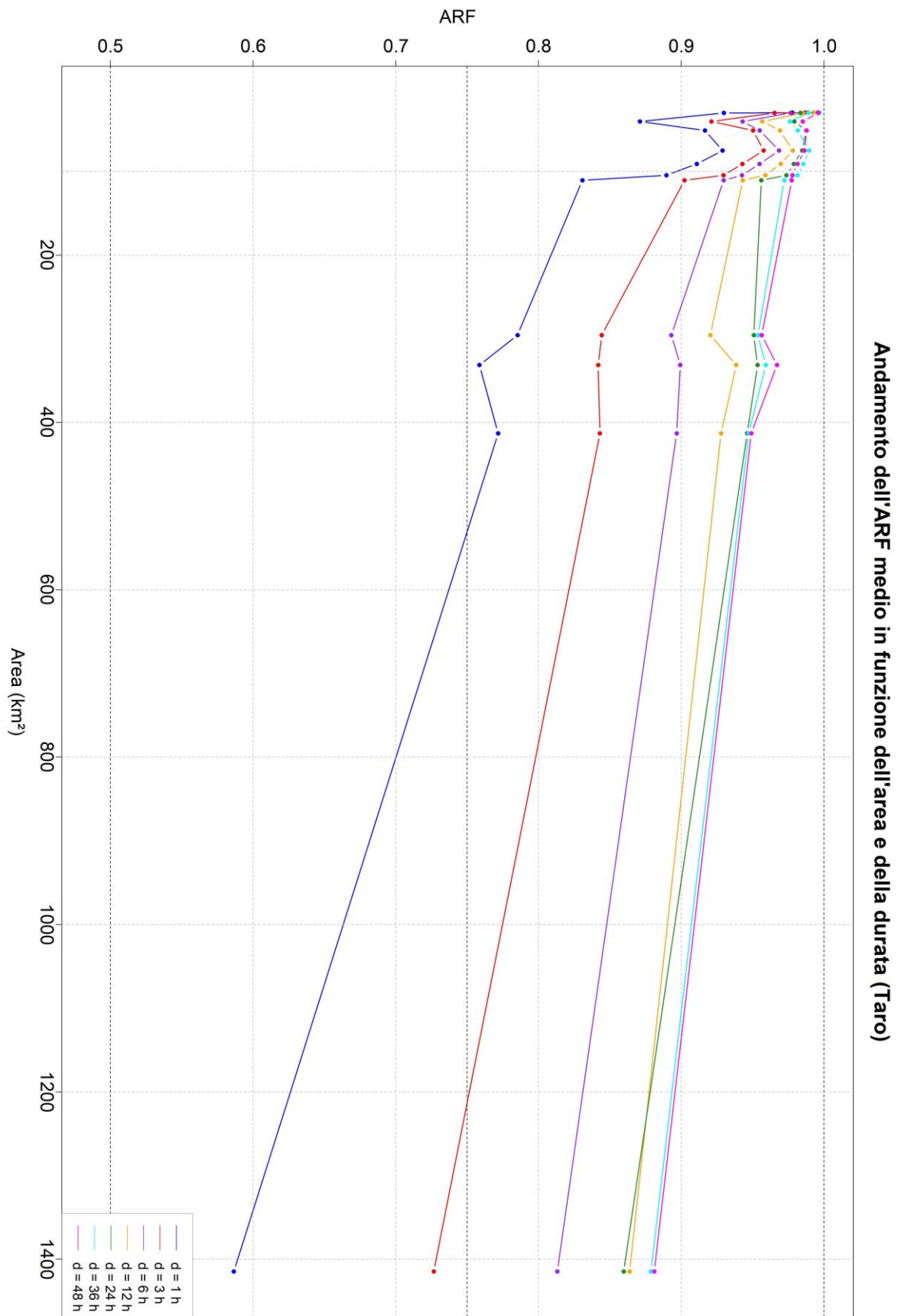
```

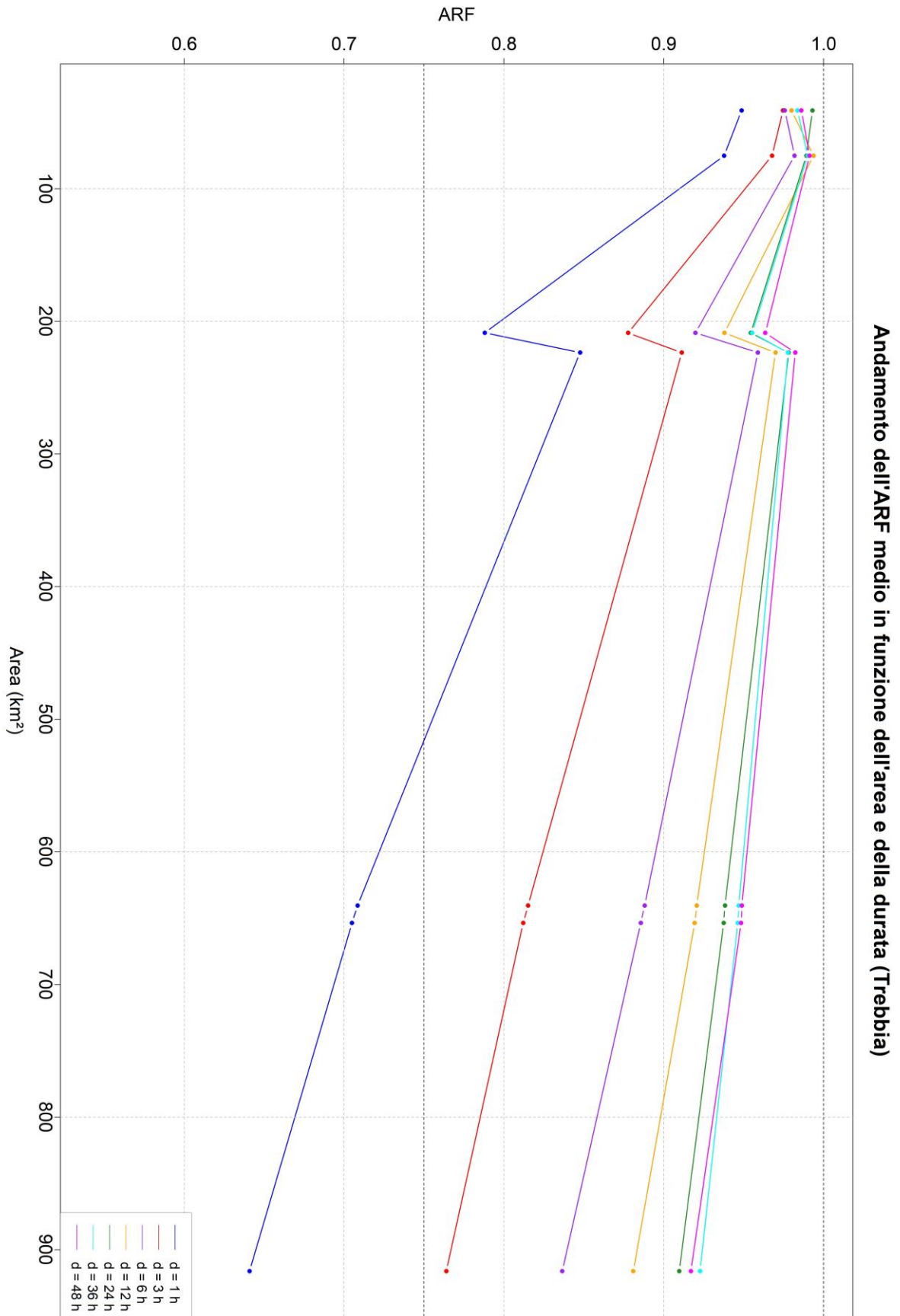
```

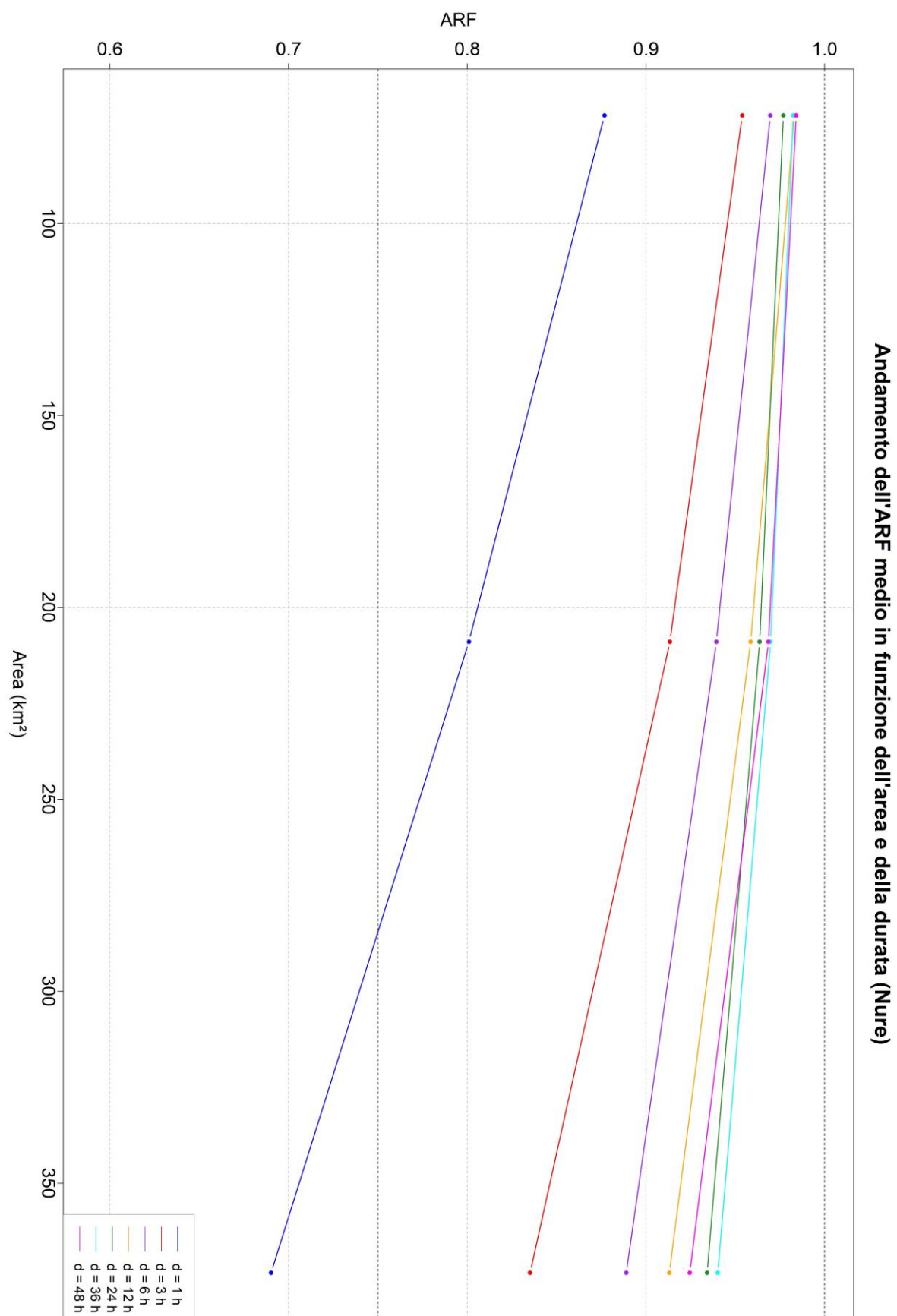
# funzione Moisello Papiri
ARF_MP <- function(parametri) {
  c1 <- parametri[1]
  c2 <- parametri[2]
  c3 <- parametri[3]
  c4 <- parametri[4]
  c5 <- parametri[5]
  for (p in 2 : ncol(ARF))
  {
    ARF[, p] <- 1 - exp(c1 * (A ^ c2) * (durata[p - 1] ^ (c3 -
exp(c4 * (A ^ c5))))))
    ARF_k <- ARF[, p]
    ARF_vec <- c(ARF_vec, ARF_k)
  }
  SSQ <- sum((ARF_oss_vec - ARF_vec)^ 2)
  N <- length(ARF_oss_vec)
  NSE <- (1 - SSQ / (var(ARF_oss_vec) * (N - 1)))
  p <- length(parametri)
  NSEadj <- 1 - (N - 1) / (N - p) * (1 - NSE)
  R2 <- cor(ARF_oss_vec, ARF_vec)^ 2
  R2adj <- 1 - (N - 1) / (N - p) * (1 - R2)
  ratio <- NULL
  for (t in 2 : ncol(ARF)) {
    ARF_mod <- ARF[, t]
    ARF_sper <- ARF_oss[, t]
    ratio_t <- sum((ARF_mod - ARF_sper) / ARF_sper)
    ratio <- c(ratio, ratio_t)
  }
  BIAS_rel <- (1 / length(durata)) * (1 / length(A)) *
sum(ratio)
  OUTPUT <- list(SSQ, NSE, NSEadj, R2, R2adj, BIAS_rel)
  return(OUTPUT)
}

```

Andamento dell'ARF medio in funzione dell'area e della durata







Boxplot dell'ARF annuale

Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 1 h (Taro)

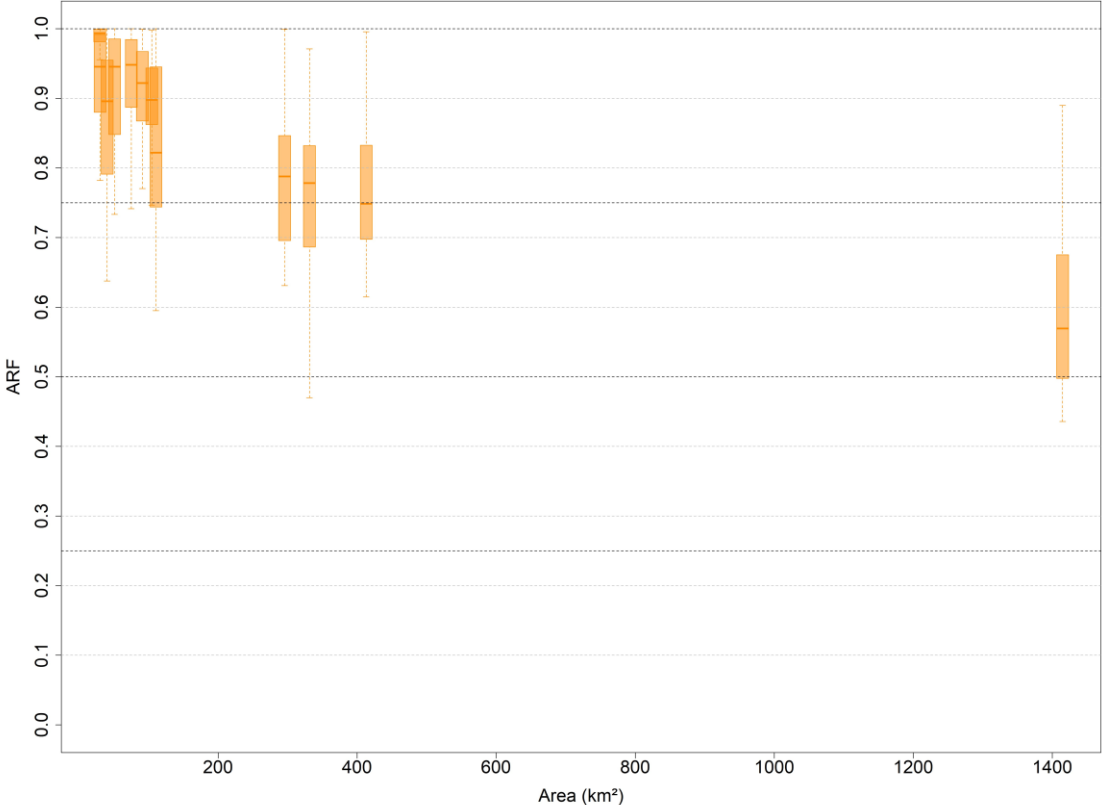


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 3 h (Taro)

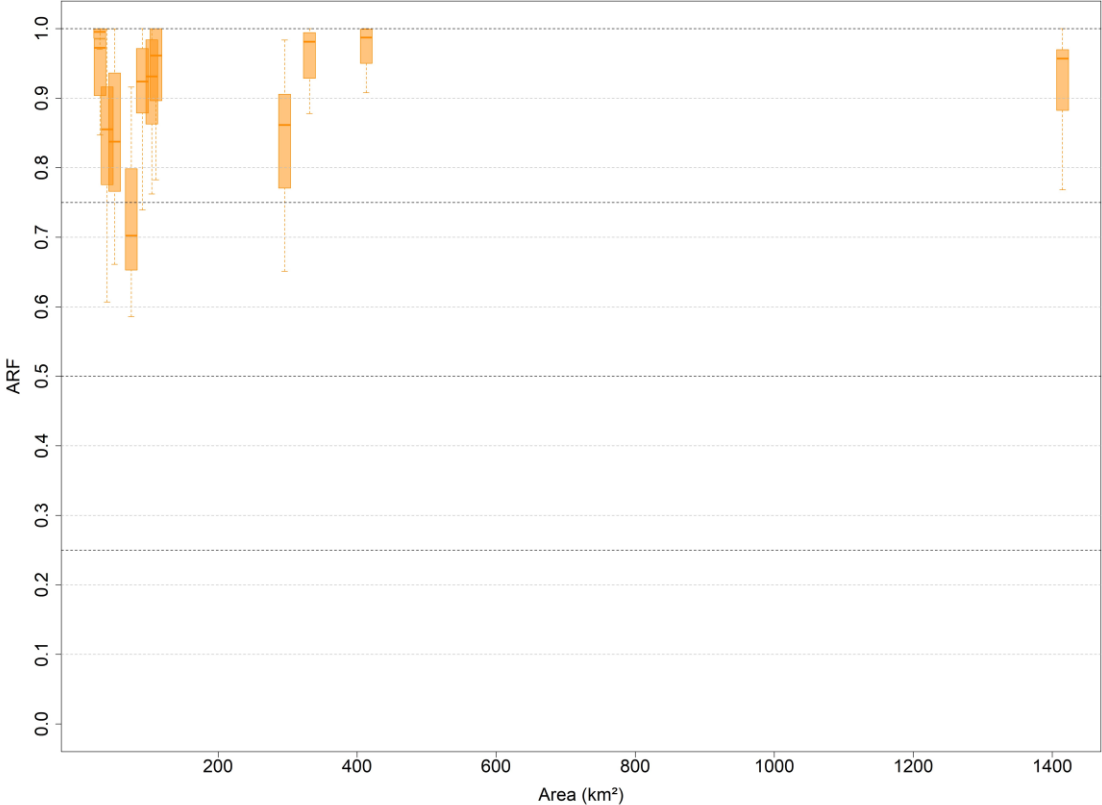


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 6 h (Taro)

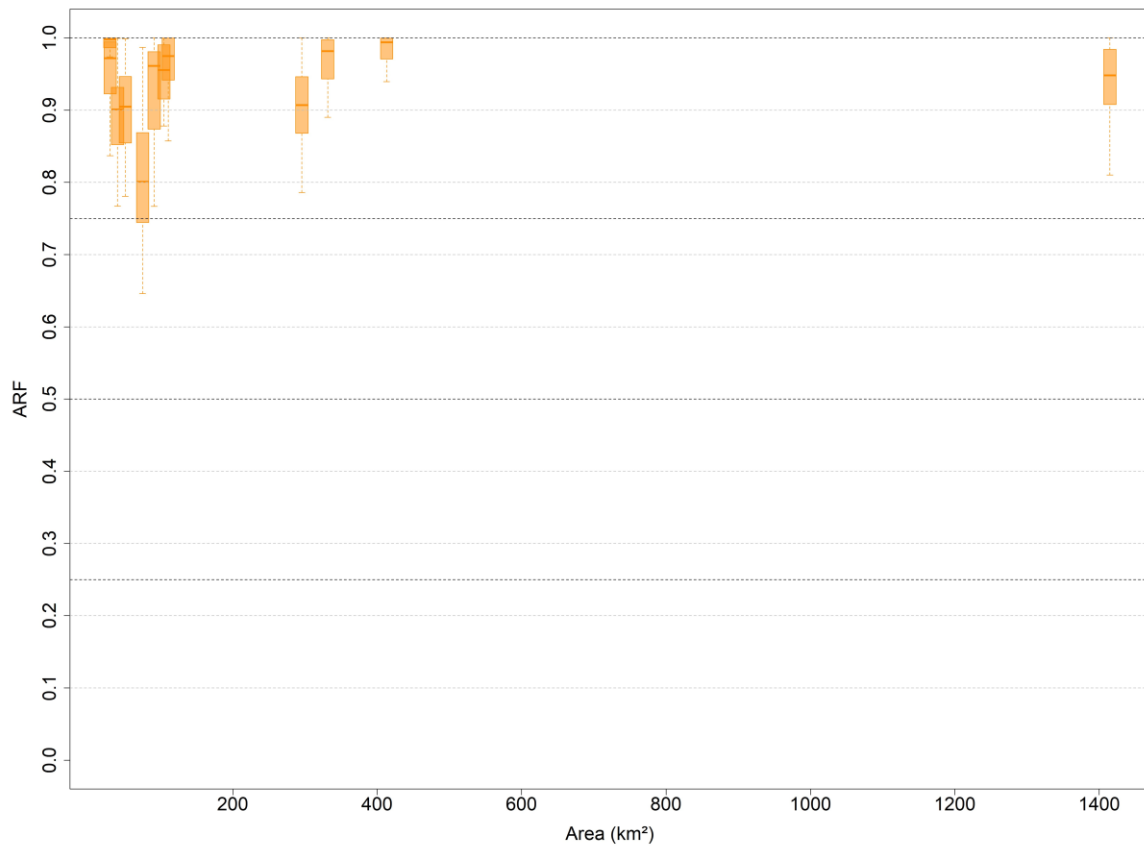


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 12 h (Taro)

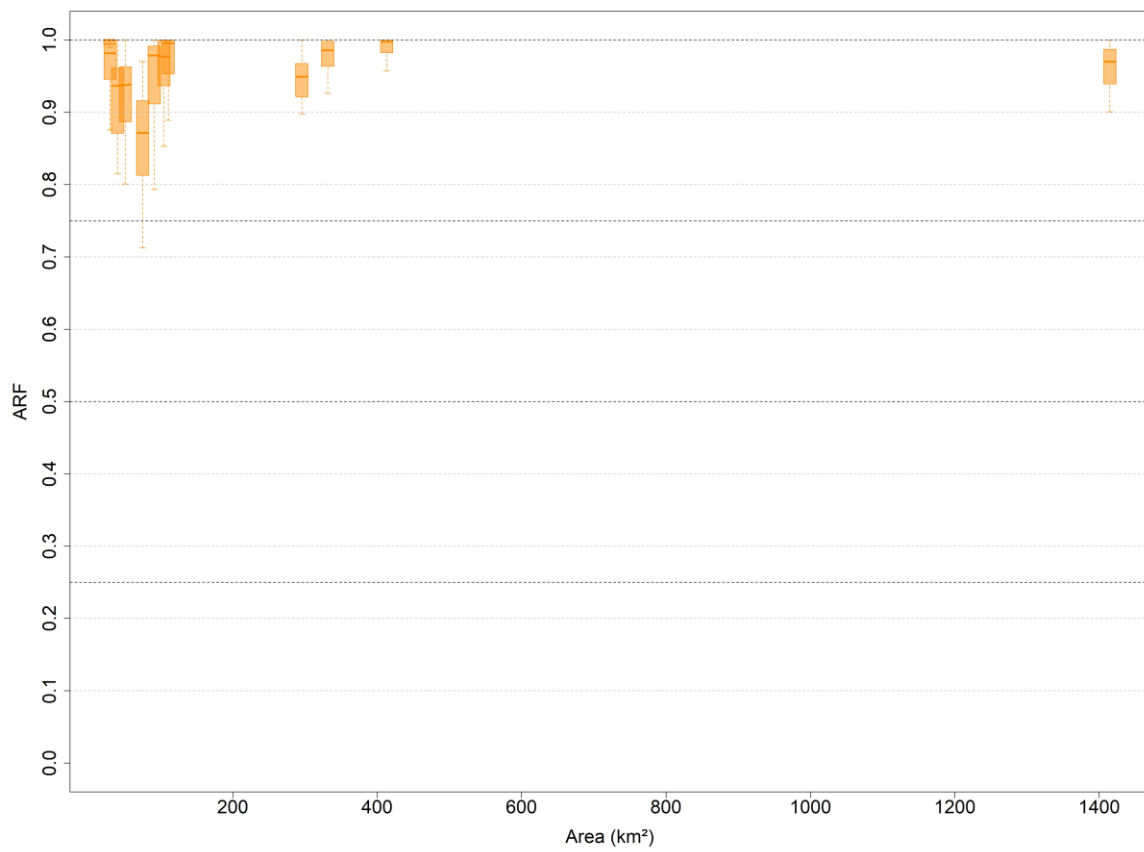


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 24 h (Taro)

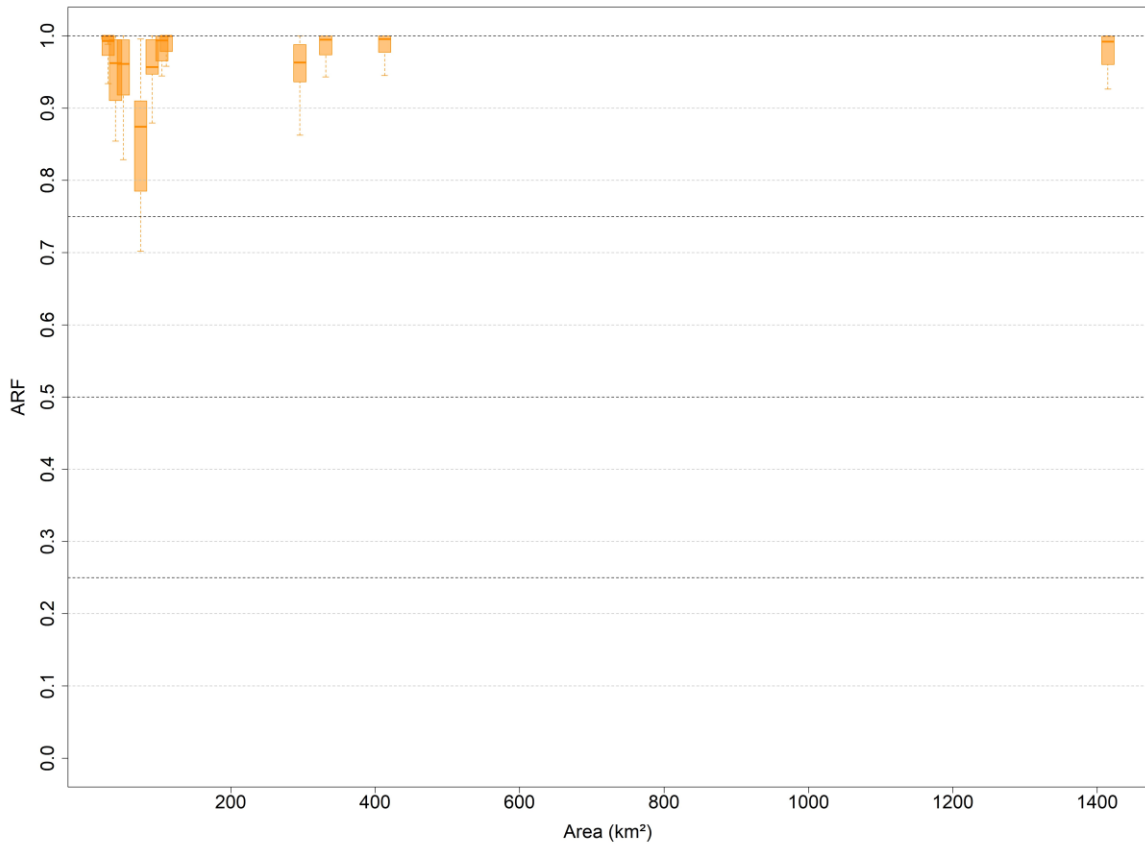


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 36 h (Taro)

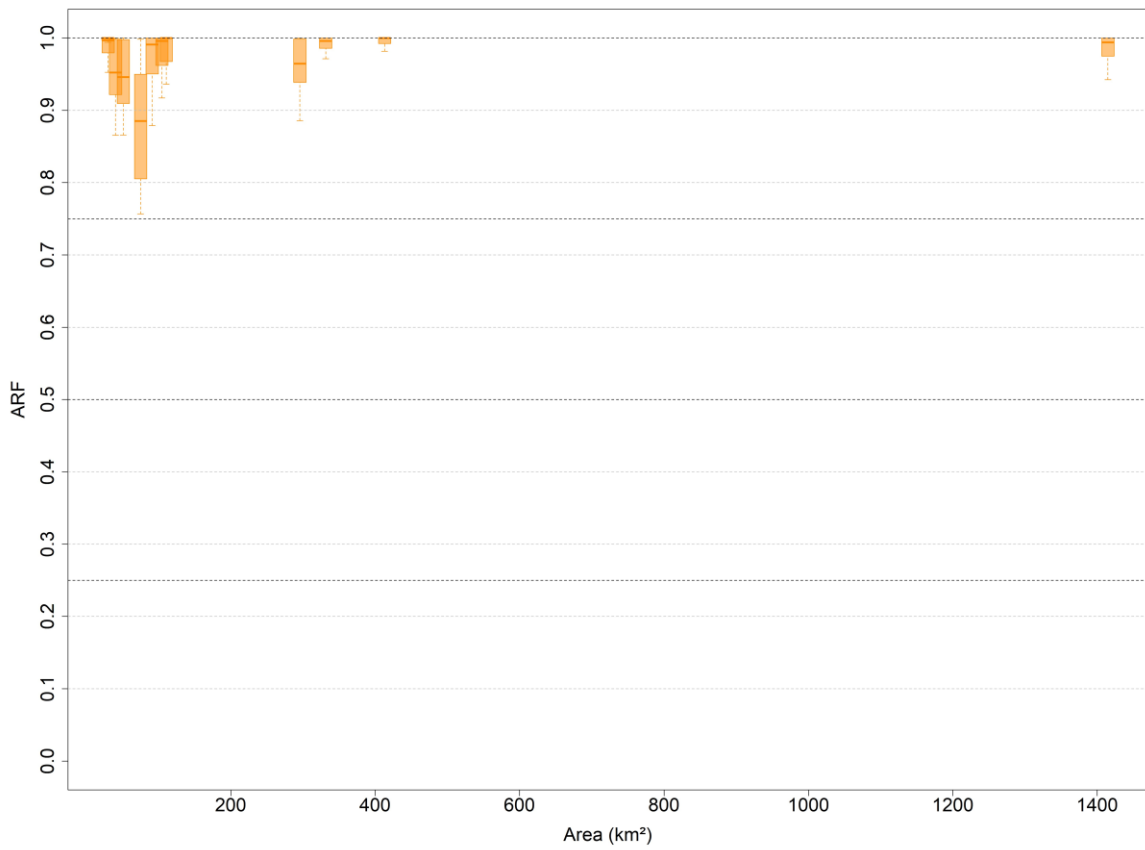


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 48 h (Taro)

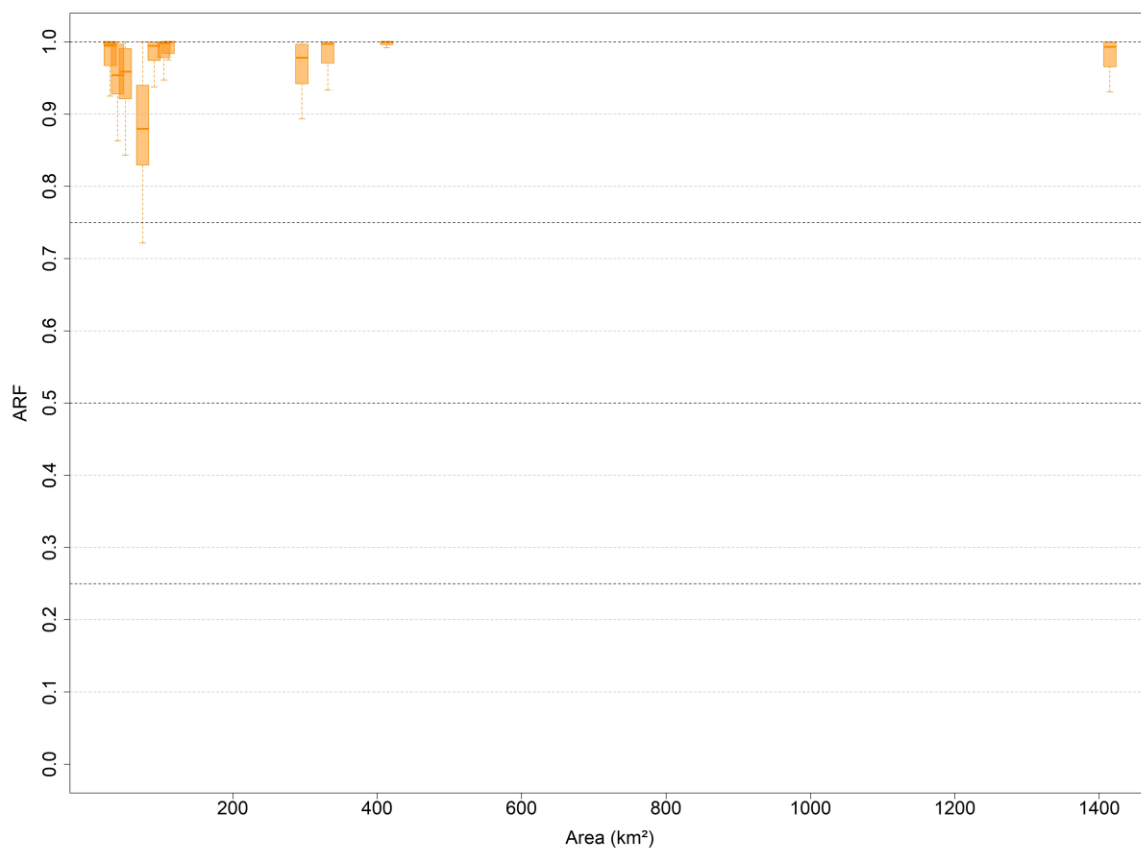


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 1 h (Trebbia)

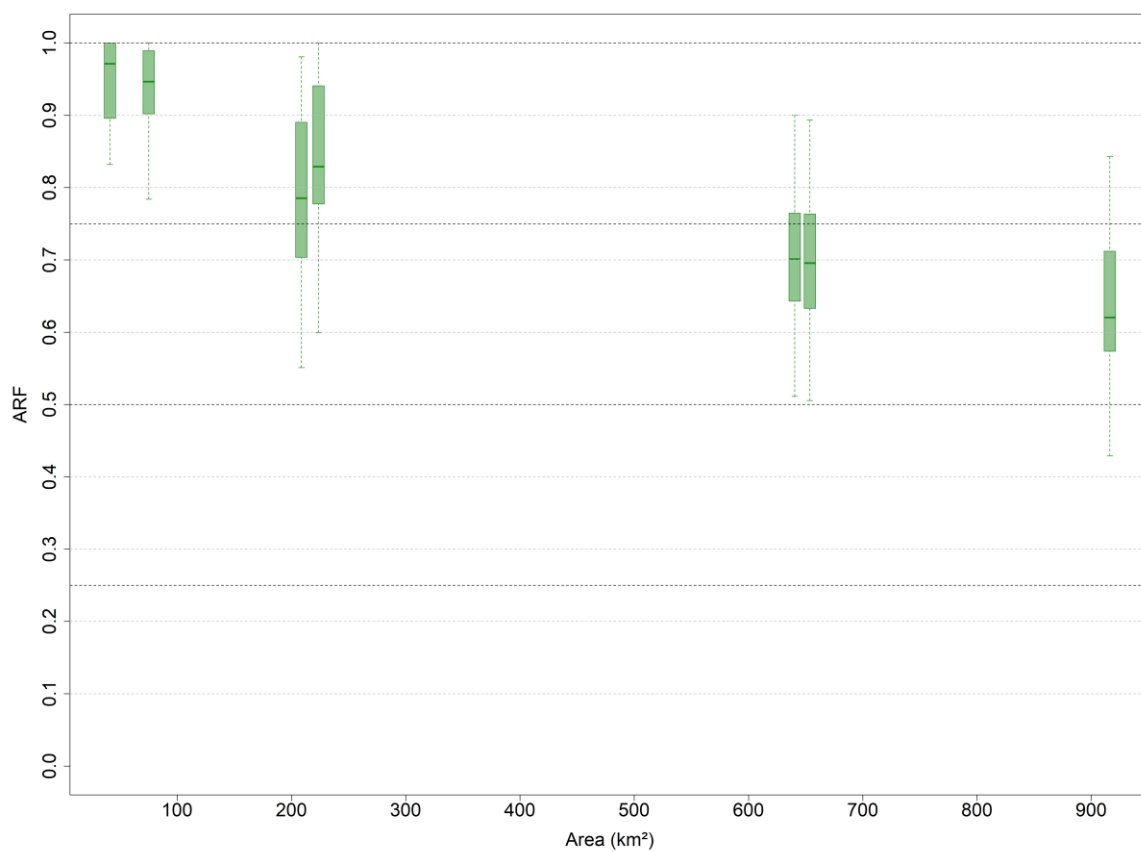


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 3 h (Trebbia)

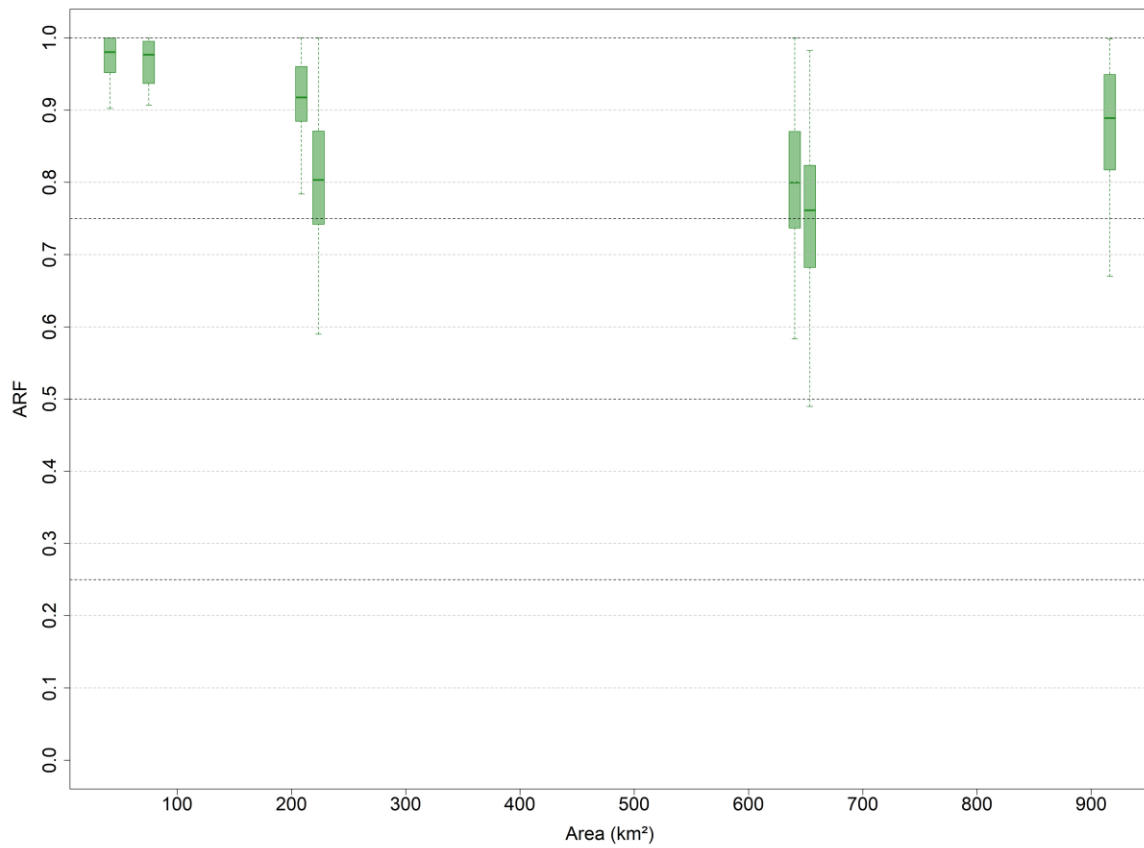


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 6 h (Trebbia)

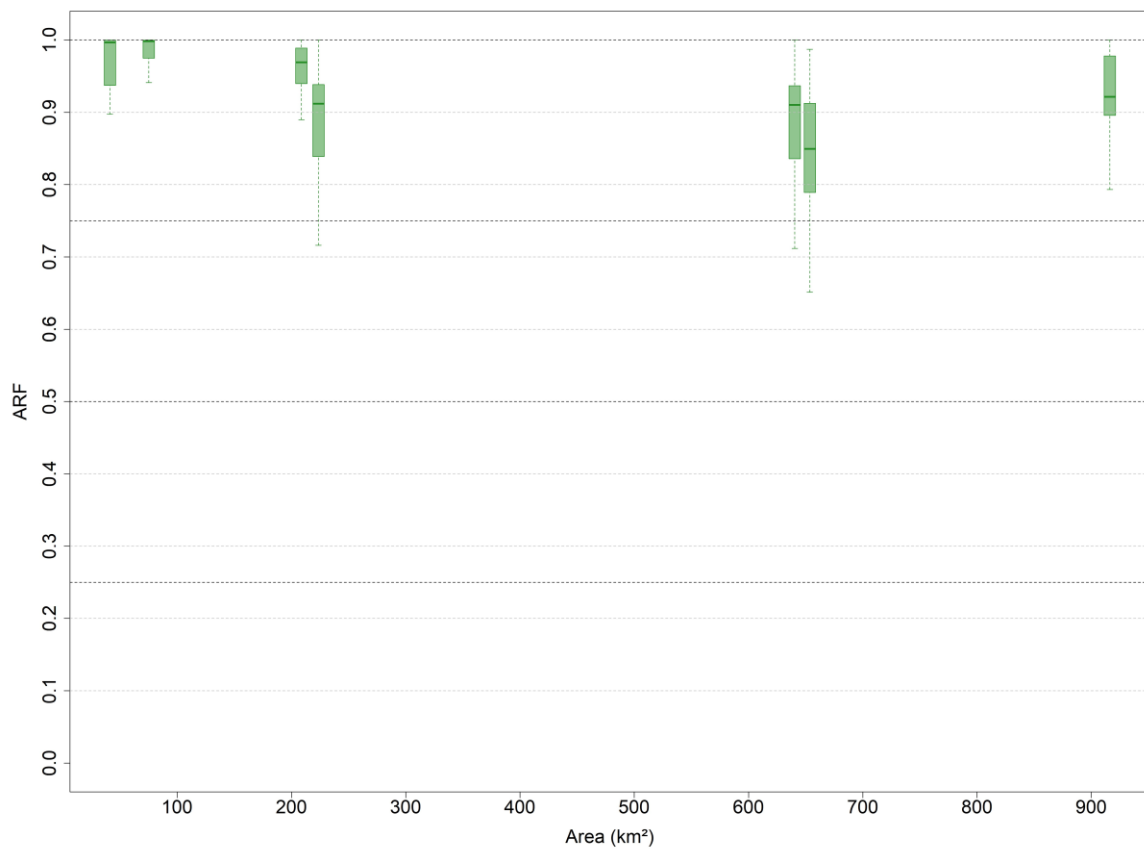


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 12 h (Tebbia)

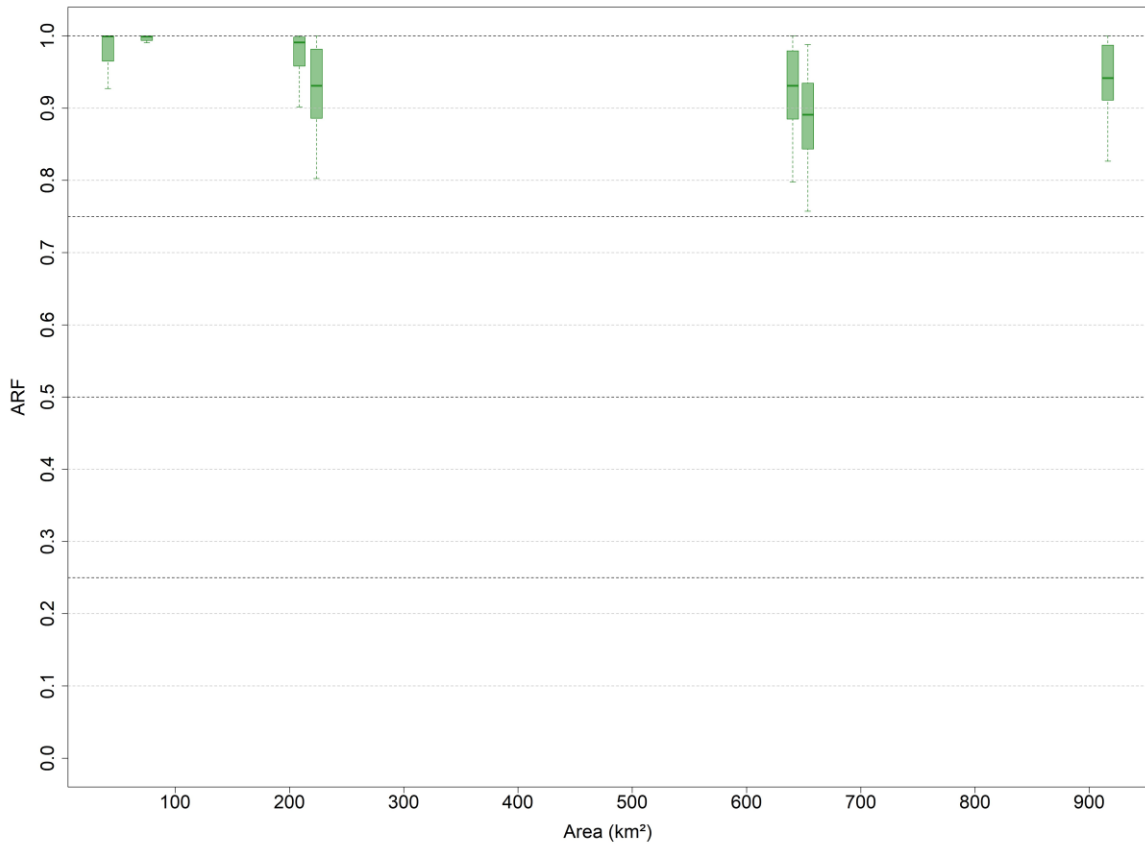


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 24 h (Tebbia)

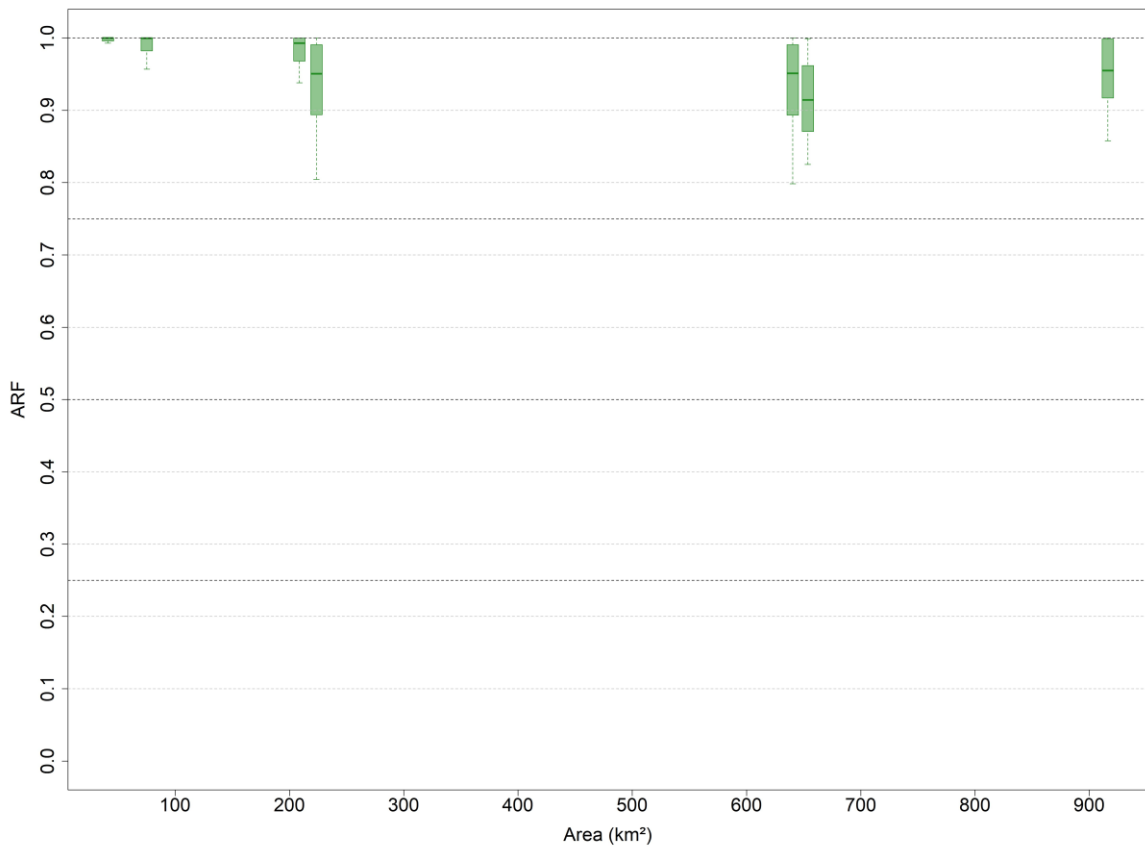


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 36 h (Trebbia)

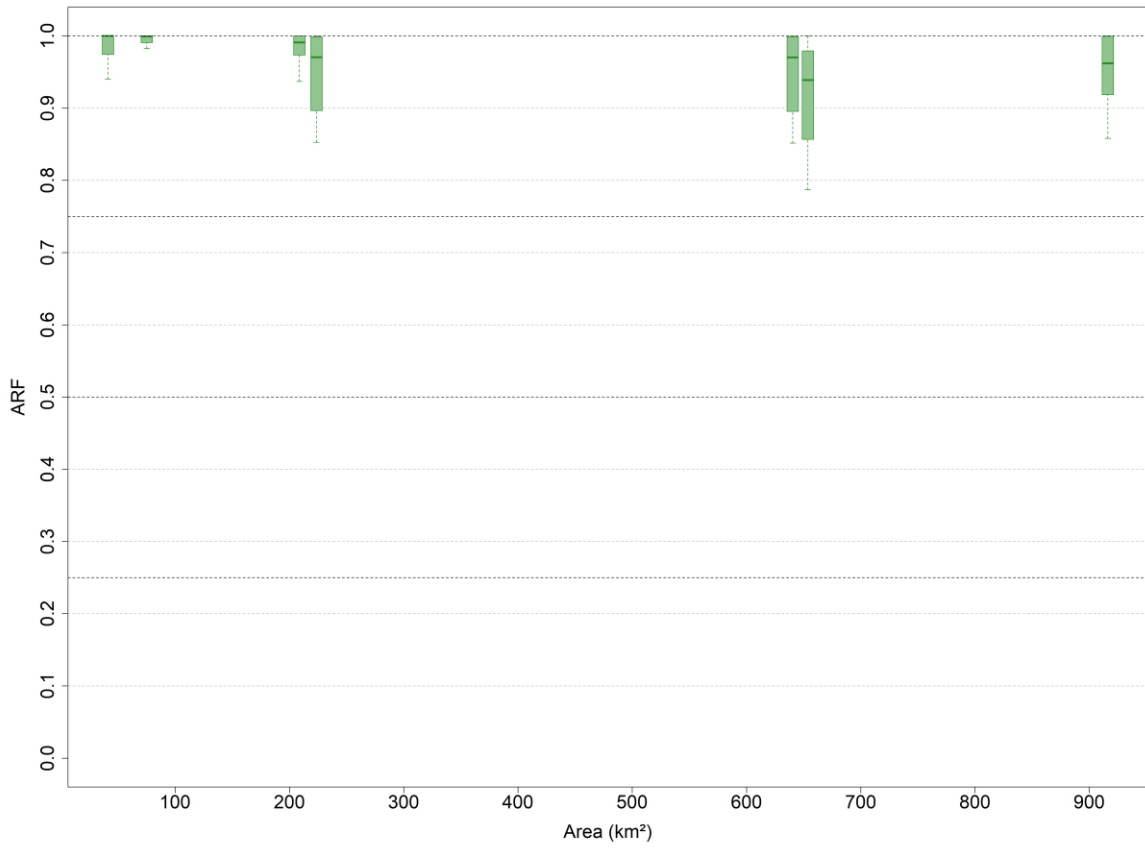


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 48 h (Trebbia)

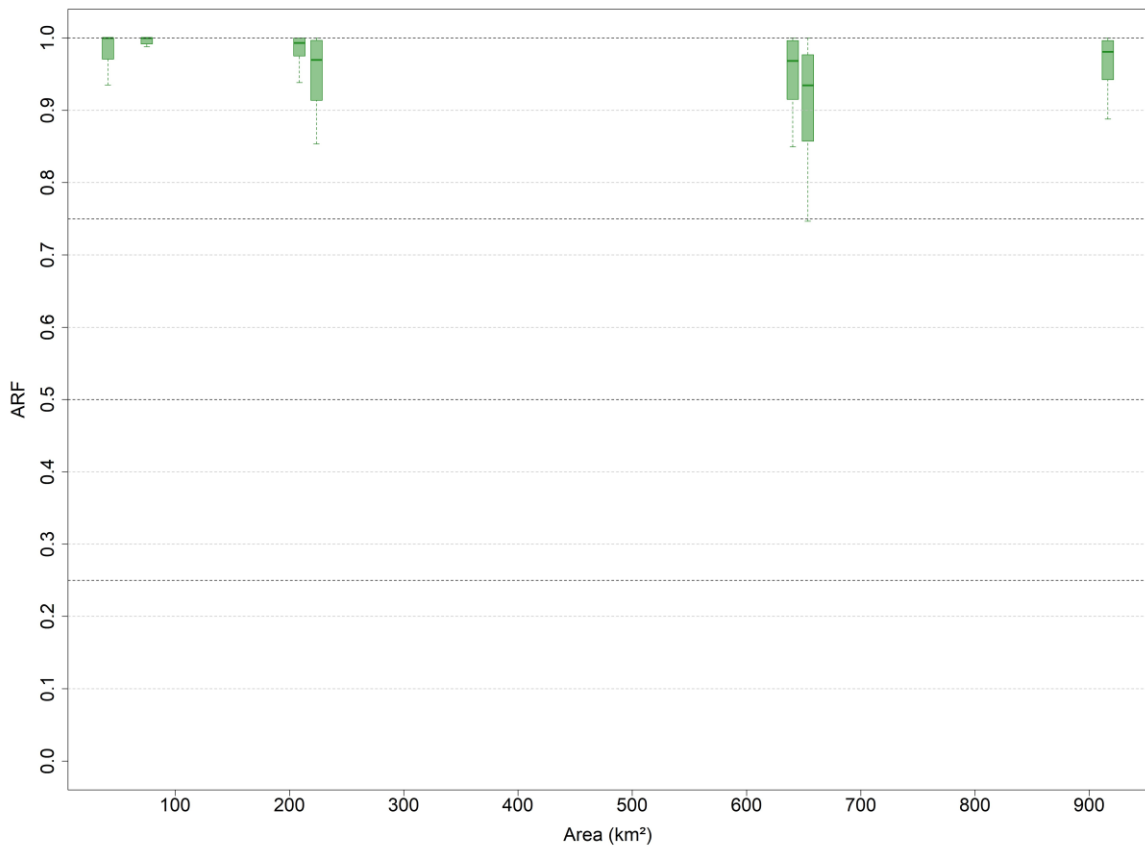


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 1 h (Nure)

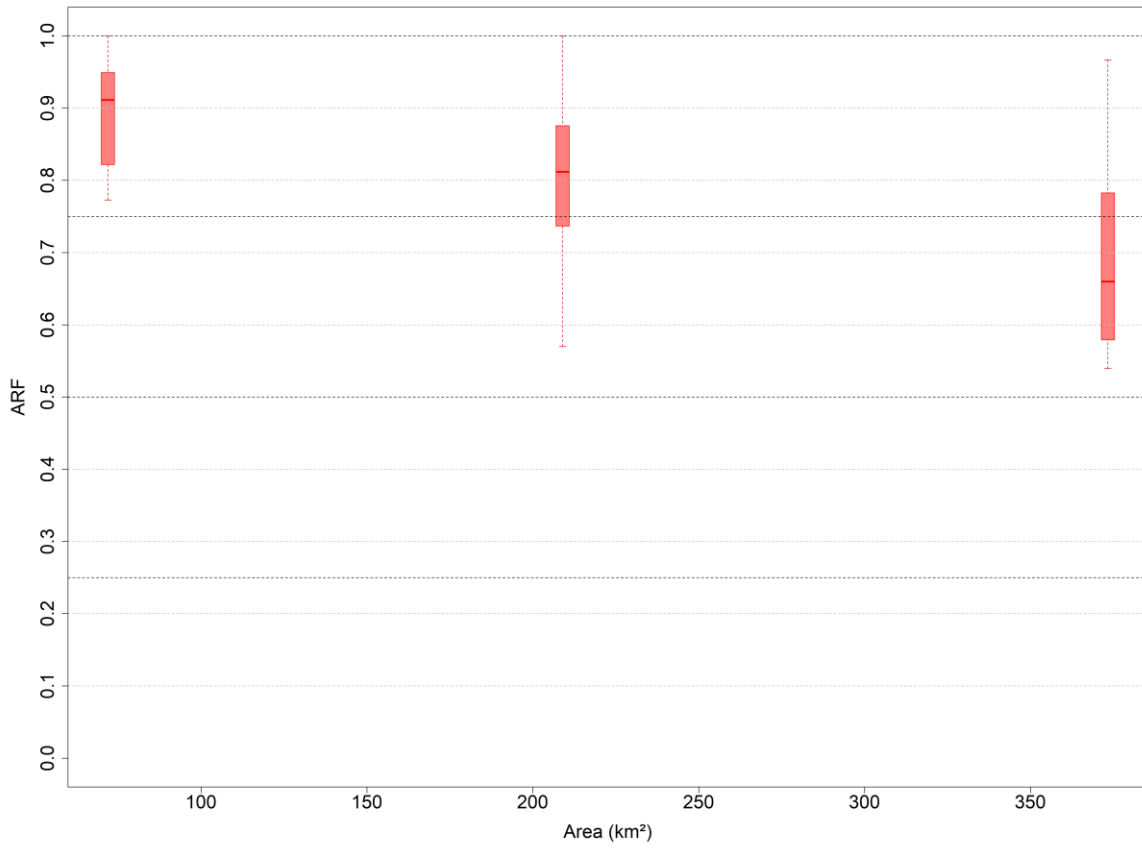


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 3 h (Nure)

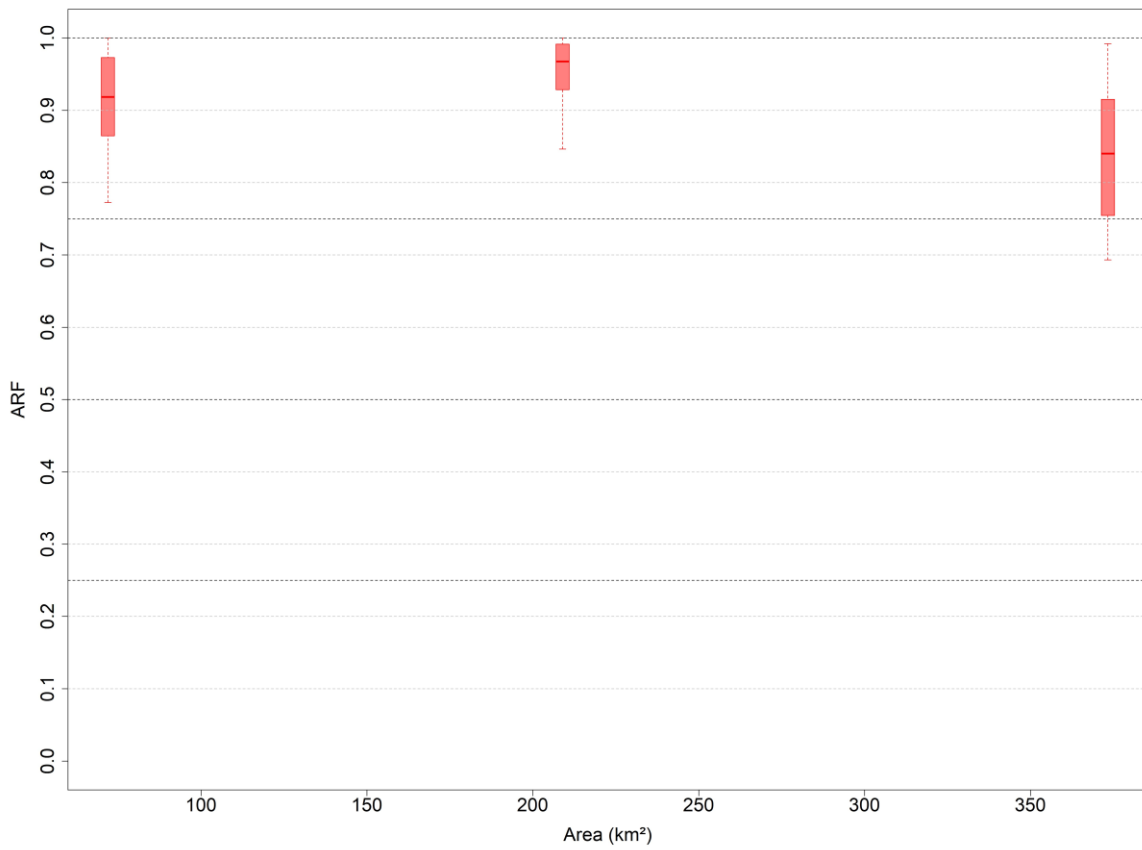


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 6 h (Nure)

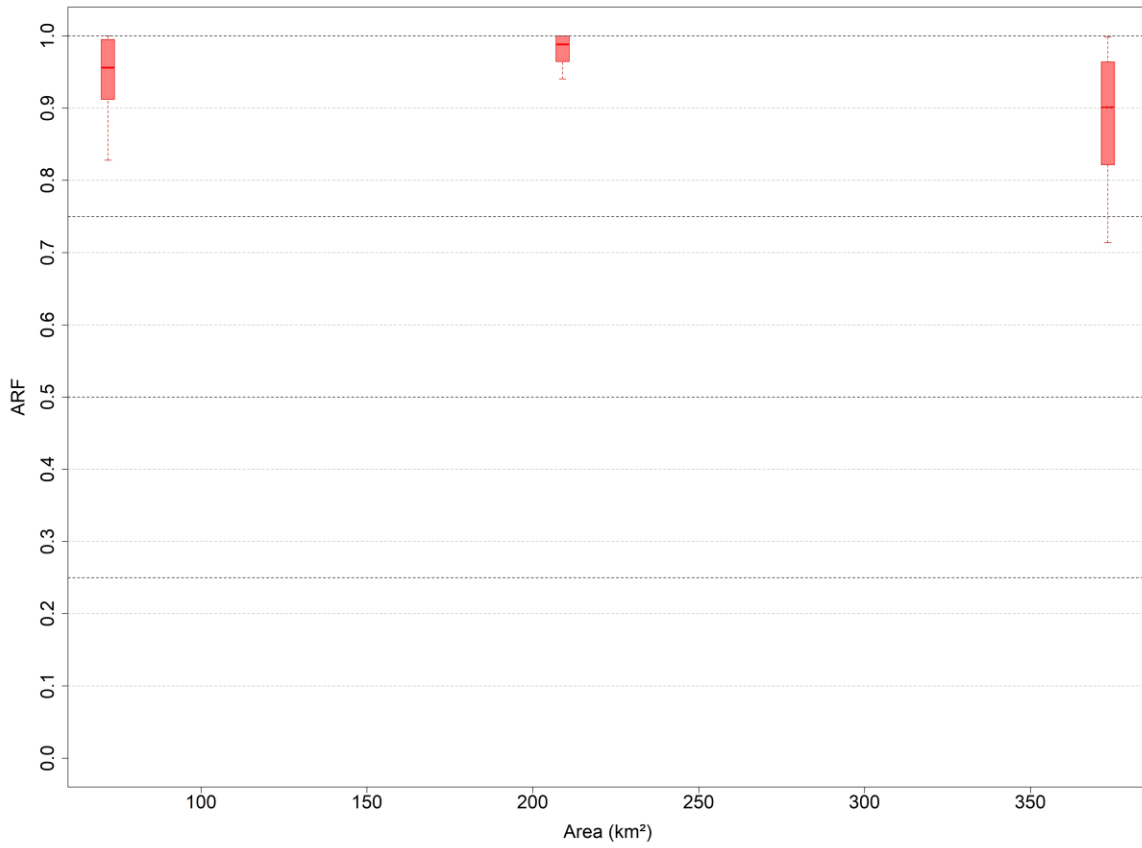


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 12 h (Nure)

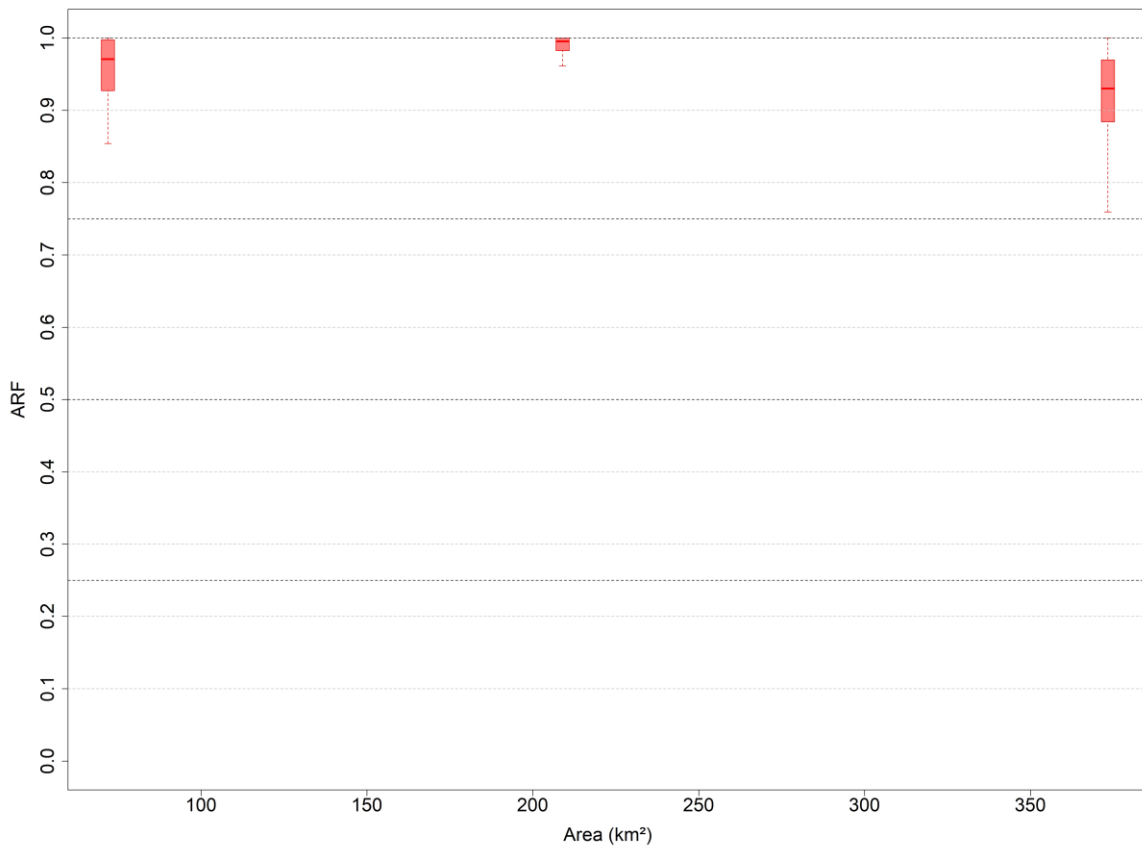


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 24 h (Nure)

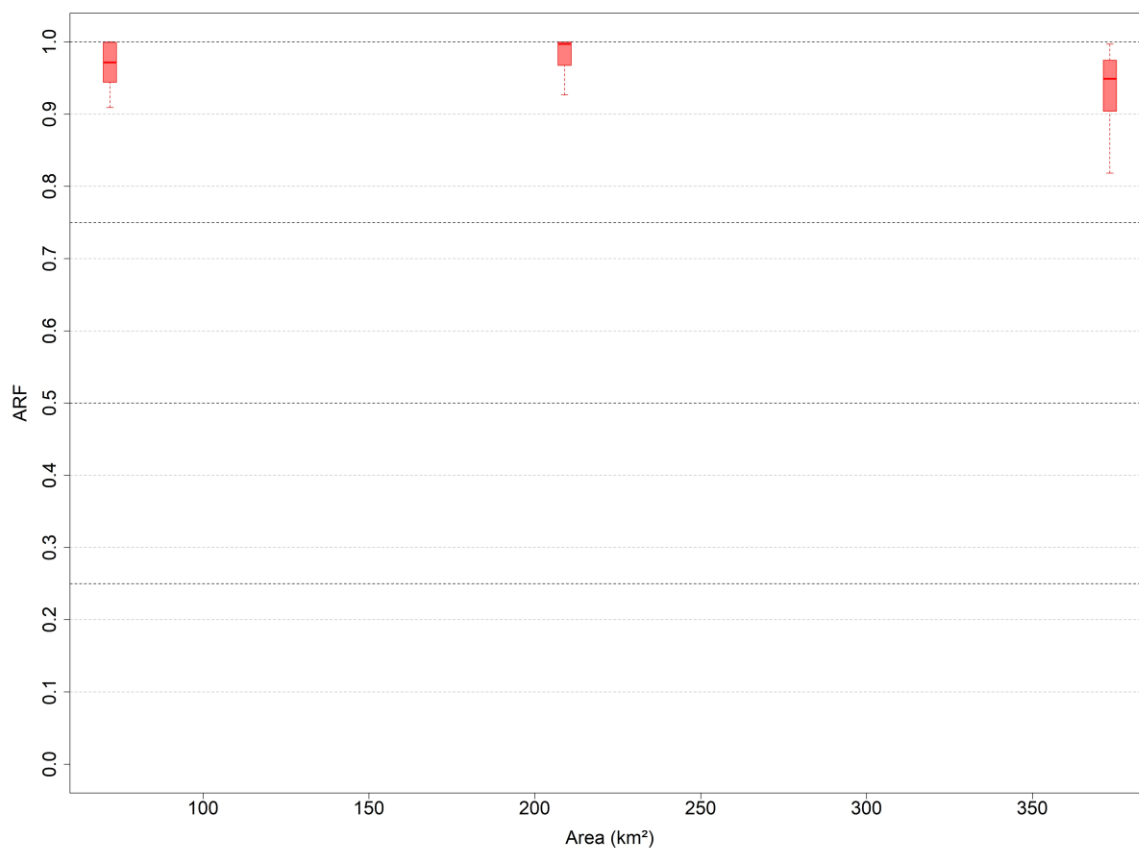


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 36 h (Nure)

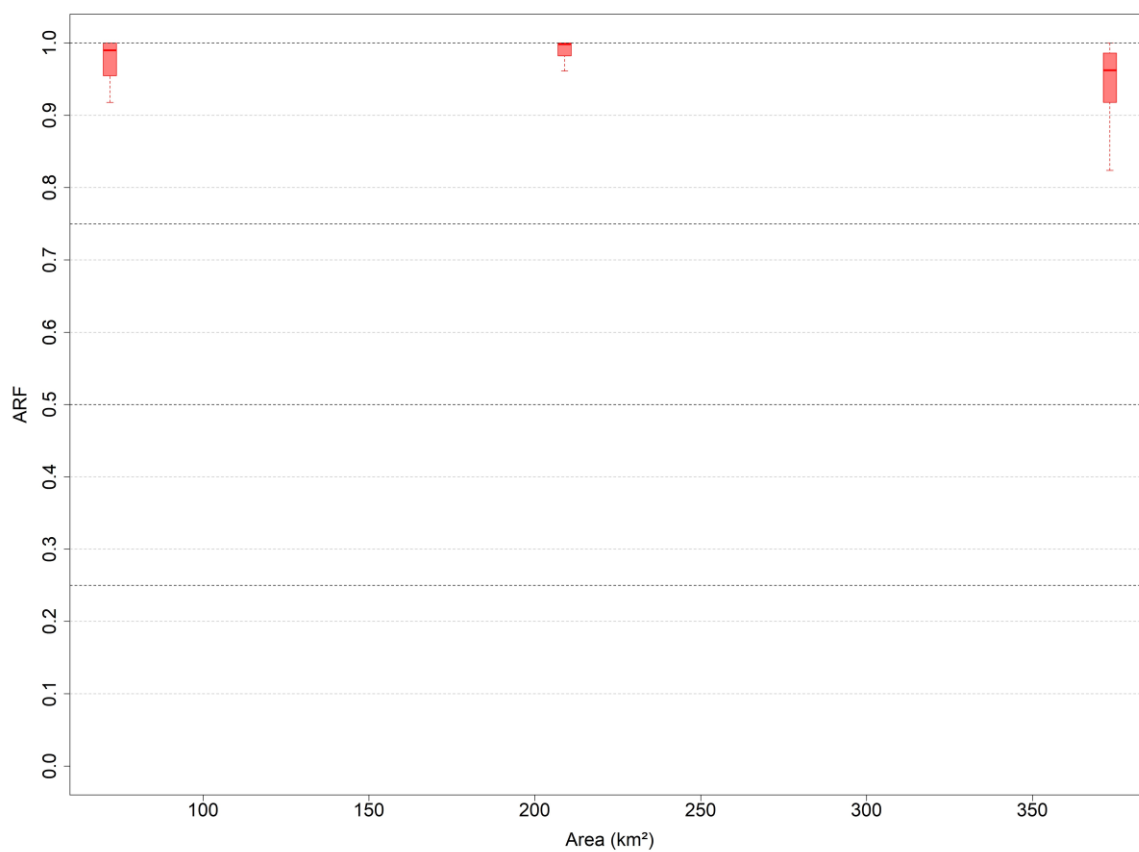
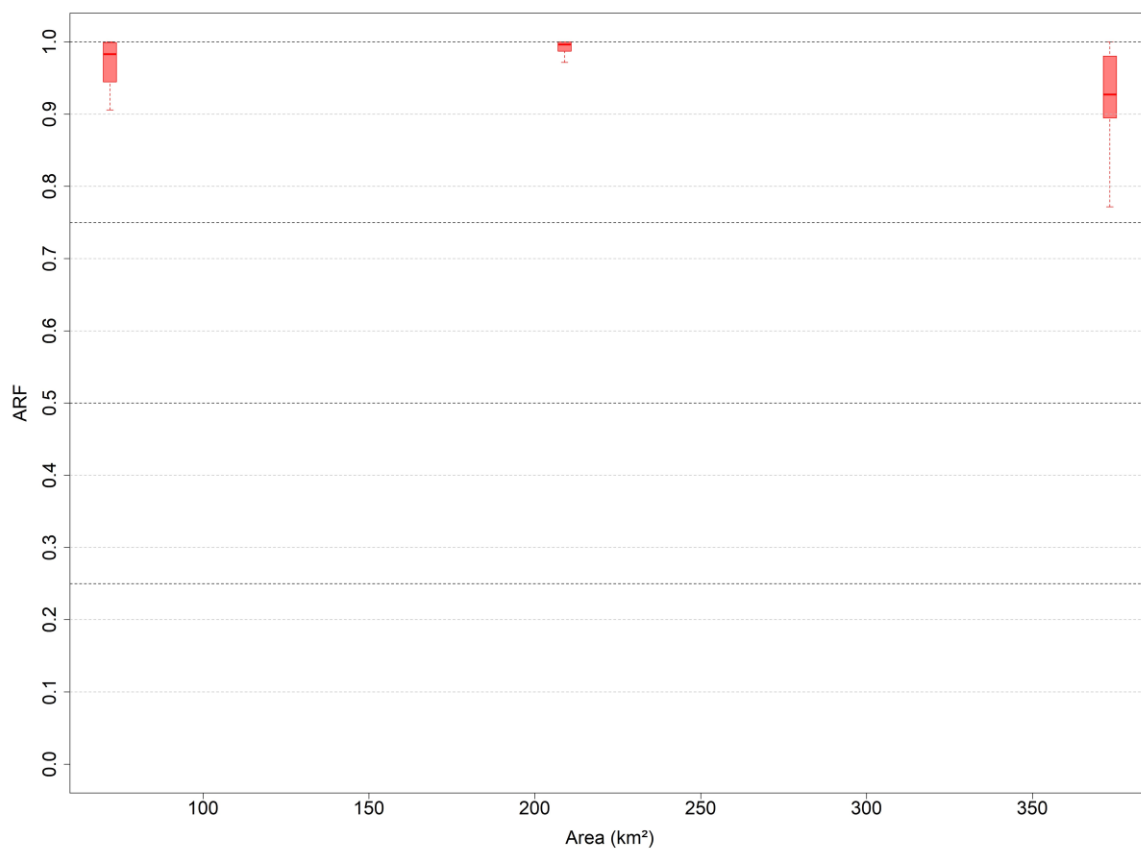


Diagramma a scatola e baffi dell'ARF annuale in funzione dell'area per una durata fissata di 48 h (Nure)



Bibliografia

- F. Archi, A. Botto, G. Braca, M. Bussettini, M. Casaioli, T. Forte, B. Lastoria, G. Marchetti, S. Mariani, C. Peruzzi, F. Piva, R. Tropeano, C. Vendetti. “Bilancio idrologico nazionale: focus su siccità e disponibilità naturale della risorsa idrica rinnovabile. Aggiornamento al 2022”. ISPRA, Rapporti 388/2023, Roma, 2023.
- G. Antolini, L. Auteri, V. Pavan, F. Tomei, R. Tomozeiua, V. Marletto. “A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961-2010”. *International Journal of Climatology*, 36, agosto 2015.
- E. Ferrari, S. Gabriele, F. Rossi, P. Versace. “Valutazione delle piene in Calabria”. CNR-IRPI, Geodata, 1989.
- F. Pascuzzi, Catalano E. "Stima della pioggia areale: alcune applicazioni sul bacino del Crati". CNR-IRPI, Rapp. int. n.267, Cosenza, 1989.
- A. Penta. “Distribuzione di probabilità del massimo annuale dell'altezza di pioggia giornaliera su un bacino”. Atti XIV Convegno di Idraulica e Costruzioni Idrauliche, Napoli, 1974.
- U. Moisello, S. Papiri. “Relazione tra altezza di pioggia puntuale e raggugliata”. Atti del XX Convegno di Idraulica e Costruzioni Idrauliche, Padova, 8-10 settembre 1986.
- F. C. Bell. “The Areal Reduction Factor in Rainfall Frequency Estimation”. Institute of Hydrology Report No 35, Natural Environment Research Council, UK, 1976.
- US Weather Bureau. “Rainfall Intensity-Frequency Regime: Part 1 - The Ohio Valley”. Technical Paper No. 29, US Weather Bureau, Washington, D.C, 1957.
- P. S. Eagleson. “Dynamics of flood frequency”. *Water Resour. Res.* 8, 878–898, 1972.G.
- Supino. “Le reti idrauliche”. Patron, Bologna, 1964.
- C. Mineo, E. Ridolfi, F. Napolitano, F. Russo. “The areal reduction factor: A new analytical expression for the Lazio Region in central Italy”. *Journal of Hydrology*, 560, 471-479, 17 marzo 2018.
- S. Kao, S. DeNeale, E. Yegorova, J. Kanney, M. Carrb. “Variability of precipitation areal reduction factors in the conterminous United States”. *Journal of Hydrology*, X 9, 2020.
- K. Breinl, H. Müller-Thomy, G. Blöschl. “Space–Time Characteristics of Areal Reduction Factors and Rainfall Processes”. *Journal of Hydrometeorology*, 21, 671-687, aprile 2020.

- B. Bacchi, R. Ranzi. "On the derivation of the areal reduction factor of storms". *Atmospheric Research*, 42, 123-135, 1996.
- C. Svensson, D. A. Jones. "Review of methods for deriving areal reduction factors". *Journal of Flood Risk Management*, 3, 232-245, 2010.
- G. Antolini, V. Marletto, V. Pavan, R. Tomozeiu. "Come cambia il clima in Emilia-Romagna". *Ecoscienza*, 1, 2017.
- G. Antolini, V. Marletto, V. Pavan, R. Tomozeiu. "Il progetto Eraclito-Atlante idroclimatico della regione Emilia-Romagna: i primi risultati". Conferenza dell'Associazione Italiana di AgroMeteorologia, gennaio 2009.
- V. Marletto, "Reti meteorologiche per l'applicazione della difesa integrata obbligatoria". Giornata di studio PAN, 25 settembre 2015.
- G. Antolini, V. Marletto, V. Pavan, F. Tomei, R. Tomozeiu. "Atlante idroclimatico dell'Emilia-Romagna 1961-2008". *I quaderni di Arpae*, 2010.
- G. Antolini, V. Pavan, R. Tomozeiu, V. Marletto. "Atlante climatico dell'Emilia-Romagna 1961-2015". 2017.
- S. Marchesi, C. Marsigli, V. Pavan, A. Selvini, R. Tomozeiu. "Controllo di qualità dei dati giornalieri di temperatura minima e massima e di precipitazione". *Quaderno tecnico ARPA-SIM*, ARPAE, 15, 2003.
- F. Desiato, G. Fioravanti, P. Frascchetti, W. Perconti, E. Piervitali. "Variazioni e tendenze degli estremi di temperatura e precipitazione in Italia". ISPRA, *Stato dell'Ambiente*, 37, luglio 2013.
- G. Antolini, A. Pasquali, V. Pavan, A. Vecchi, M. Tartaro, L. Angelo, G. Ricciardi, F. Tugnoli. "Bollettino mensile, a cura della Struttura Idro-Meteo-Clima". Arpae, Struttura Idro-Meteo-Clima, Anno IV, n.5, maggio 2023.
- C. Cardinali, R. Foraci, M. S. Tesini, M. Pizziolo. "Rapporto degli eventi meteorologici di piena e di frana dell'1-4 maggio 2023". Arpa Emilia Romagna Struttura Idro Meteo Clima, 3 giugno 2023.
- E. Covi, R. Stanzani. "Rapporto dell'evento meteorologico dal 12 al 14 maggio 2023". Arpae Emilia-Romagna - Struttura Idro-Meteo-Clima, 5 giugno 2023.

A. Fornasiero, M. S. Tesini, R. Foraci, L. G. Biolchi, A. Valentini, V. Pavan, G. Antolini, M. Pizziolo, M. Generali, M. Scaroni, L. Perini, J. Lelli. “Rapporto degli eventi meteorologici di piena e di frana del 16-18 maggio 2023”. Arpae Emilia-Romagna - Struttura Idro-Meteo-Clima, 7 luglio 2023.

“Nuovi prodotti meteo-climatici su open data”. Ambiente Regione Emilia–Romagna, 30 giugno 2022, ambiente.regione.emilia-romagna.it/it/cambiamenti-climatici/notizie/attualita-1/2022/nuovi-prodotti-meteo-climatici-su-open-data-1.

G. Antolini. “ERG5_Eraclito - Dataset climatico dal 1961”. Dati Arpae Emilia–Romagna, 31 ottobre 2019, dati.arpae.it/dataset/erg5-interpolazione-su-griglia-di-dati-meteo.

G. Antolini. “ERG5 - Dataset meteo orario e giornaliero dal 2001”. Dati Arpae Emilia–Romagna, 13 giugno 2017, dati.arpae.it/dataset/erg5-eraclito.

“Rete di monitoraggio idrometeorologica”. Arpae Emilia-Romagna, pubblicato il 2 ottobre 2020, modificato il 21 novembre 2021, www.arpae.it/it/temi-ambientali/meteo/dati-e-osservazioni/rete-di-monitoraggio-idrometeorologica.

“Clima passato, presente e futuro”: il 2° episodio del podcast Agrifuturo”. Food&tec_, 3 aprile 2023, www.foodandtec.com/n/clima-passato-presente-e-futuro-il-2-episodio-del-podcast-agrifuturo.

“Alluvioni in Emilia-Romagna e Marche”. Climate Media Center Italia, 18 maggio 2023, <https://www.climatemediacenteritalia.it/alluvioni-emilia-romagna-e-marche/>.

“Le alluvioni di maggio 2023: un’analisi scientifica”. CIMA Research Foundation, 19 maggio 2023, <https://www.cimafoundation.org/news/le-alluvioni-di-maggio-2023-una-analisi-scientifica/>.

“Alluvione in Emilia-Romagna: l’anomalia combinata di cielo e terra”. CIMA Research Foundation, 5 maggio 2023, <https://www.cimafoundation.org/news/alluvione-in-emilia-romagna-lanomalia-combinata-di-cielo-e-terra/>.

“Alluvioni di maggio 2023 in Emilia orientale e Romagna”. Nimbus, 20 maggio 2023, <http://www.nimbus.it/eventi/2023/230519AlluvioneEmiliaRomagna.htm>.

E. Ferri. “Variabilità spaziale e temporale degli estremi di precipitazione sub-giornalieri osservati in Emilia-Romagna”. Tesi di laurea magistrale, Università degli Studi di Bologna, A.A. 2017-2018.

C. Bonapace. “La valutazione sperimentale del fattore di riduzione delle piogge all’area”. Tesi di laurea, Università degli Studi di Bologna, A.A. 1999-2000.

L. Auteri. “Analisi climatica ad alta risoluzione delle precipitazioni sul nord Italia (1961-2005)”. Tesi di dottorato, Università degli Studi di Bologna, 2013.

A. Langousis. “The Areal Reduction Factor (ARF) a Multifractal Analysis”. Tesi di Master, Massachusetts Institute of Technology, 2005.

L. Comandini. “Analisi statistiche di dati idro-meteorologici del bacino del fiume Santerno (BO) nell’ultimo secolo”. Tesi di laurea, Università degli Studi di Bologna, A.A. 2014-2015.

F. Ambrosiani. “Interpolazione di mappe DEM mediante algoritmi di distanza inversa pesata e funzioni a base radiale”. Tesi di laurea magistrale, Politecnico di Milano, A.A. 2012-2013.

A. Fecondini. “Analisi di frequenza congiunta degli estremi pluviometrici e idrometrici nel bacino del Reno”. Tesi di laurea magistrale, Università degli Studi di Bologna, A.A. 2006-2007.

