

Politecnico di Torino

Corso di Laurea Magistrale in
Ingegneria per l'Ambiente e il Territorio



Tesi di laurea

*Confronto di due tecniche di misura della
precipitazione in Piemonte*

Candidato
Mattia Rusco

Relatore
Francesco Laio

AA 2019/20

"How can something so common and familiar be so difficult to describe?"

The water book, Alok Jha

Abstract italiano

La misura accurata della precipitazione è di fondamentale importanza per numerose applicazioni quali, ad esempio, la corretta gestione agricola e la mitigazione dei rischi naturali. La precipitazione, liquida o solida che sia, è un fenomeno ad evoluzione rapida, per la quale risulta cruciale disporre di uno strumento di misura preciso e ad elevata risoluzione temporale. Lo studio si prefigge l'obiettivo di valutare la possibilità di affiancare ad un consolidato metodo di misura della precipitazione a terra un sistema di stima in remoto. Infatti, osservare materialmente il quantitativo di pioggia caduto al suolo permette di ottenere un'elevata accuratezza del dato, ma al tempo stesso tale metodo implica la presenza di centraline meteorologiche nella zona di interesse ed elevati costi di manutenzione delle stesse. Alcuni sistemi satellitari possiedono la capacità di misurare lo stesso fenomeno senza la necessità di installare alcuno strumento in campo: tra essi si annovera l'iniziativa *Global Precipitation Measurement (GPM)*, ovvero una costellazione di satelliti associata ad un algoritmo di analisi (*Integrated Multi-satellitE Retrievals for GPM*, o *IMERG*) in grado di fornire una stima di pioggia sulla superficie con una risoluzione spaziale di $0,1^\circ$ ed un tempo minimo di acquisizione di 30 minuti. Nello studio sono stati utilizzati i dati di questa risorsa in contrapposizione a quelli provenienti una rete pluviometrica gestita da un'azienda sita a Torino di nome 3a S.r.l.

L'analisi è stata svolta all'interno del territorio Piemontese: a causa della sua orografia complessa le condizioni ambientali variano considerevolmente in uno spazio assai ridotto. È stata eseguita una valutazione tecnica degli strumenti di misura più diffusi e successivamente è stato operato un confronto servendosi di differenti aggregazioni temporali (da 1 ora a 1 mese) al fine di valutare la capacità del sistema satellitare di cogliere la presenza e la quantità di acqua precipitata.

Considerando gli anni dal 2013 al 2017, sono state prodotte diverse serie storiche per valutare la performance di IMERG, successivamente si è ricorso ad indici categorici e statistici con relative mappe tematiche per apprezzare eventuali differenze spaziali di prestazione, permettendo di riscontrare, ad esempio, una blanda correlazione con l'elevazione del terreno.

La comparazione dei dati misurati tramite centraline pluviometriche ed i dati provenienti da IMERG mostra come la stima in remoto possieda discrete potenzialità di valutazione della precipitazione, risultando capace di rilevare la precipitazione nel 70% dei casi, pur restando uno strumento con un'affidabilità non equiparabile a quella di una rete a terra, limitando le applicazioni di precisione ma consentendo l'utilizzo dei dati per analisi spaziali a livello regionale.

English abstract

High accuracy when measuring rainfall is of paramount importance in a wide range of applications, ranging from appropriate agriculture management to the mitigation of natural hazards. Precipitation (solid or liquid), often occurs with a rapid evolution, which implies the relevant gauging tools being managed with high temporal resolution.

This study aims to assess the chance to accost a traditional rain gauge network with a remote sensing system. Measuring rain directly from the ground allows one to achieve a high level of accuracy, in spite of the necessity to install a gauge on the interest area with consequent maintenance costs. Nowadays, some satellite systems allow one to estimate rain without having to be physically present on the spot: among them there is the *Global Precipitation Measurement* initiative, or *GPM*, which consists of a satellite constellation together with an associated algorithm named *Integrated Multi-satellitE Retrievals for GPM*, also known as *IMERG*, capable of estimating rainfall on the ground with a $0,1^\circ$ spatial and an half-hour temporal resolution. This particular source has been compared to data coming from a rain gauge network managed by a private company named *3a S.r.l.*, which is based in Turin.

The analysis has been carried out in Piedmont, a region in north western Italy characterized by a complex orography and proximity to the Alps, considering the years from 2013 to 2017. This project includes a preliminary study on the state-of-the-art technologies available on the market and then focuses on the comparison between the two aforementioned data sources using different time scales, going from one hour to one month, in order to evaluate the ability of the GPM to estimate rainfalls amounts.

The output includes historical series for each grid cell, graphs and maps to eventually assess differences among different areas: e.g. a slight correlation between elevation and performance has been observed.

This approach shows that the IMERG performances are fairly good, having a success rate of 70% in detecting rainfall, but IMERG lacks in accuracy about quantities, making it unlikely to be considered for precision applications, yet allowing its use for spatial analysis on a regional scale.

Sommario

1. Introduzione	1
2. Stato dell'arte per le tecnologie di misurazione della precipitazione.....	4
2.1. <i>Radar a terra</i>	6
2.2. <i>Pluviometri</i>	11
2.2.1. <i>Caratteristiche tecniche</i>	12
2.2.2. <i>Modelli disponibili sul mercato</i>	15
2.2.3. <i>Rete locale utilizzata</i>	19
2.2.4. <i>Database e controllo di qualità</i>	23
2.3. <i>Satelliti</i>	25
2.3.1. <i>GPM</i>	29
2.3.2. <i>Strumentazione GPM</i>	31
2.3.3. <i>Segmento a terra ed algoritmi di calcolo</i>	34
3. Reperibilità dati e procedure applicate	37
3.1. <i>Scelta e reperibilità dei dati</i>	37
3.2. <i>Pre-processing</i>	42
3.3. <i>Processing</i>	47
4. Risultati.....	51
5. Discussione dei risultati.....	75
6. Conclusioni.....	77
Bibliografia	79
Appendice	83
<i>Pre-processing script</i>	96
<i>Analisi statistica</i>	100
Ringraziamenti.....	109

Indice delle figure

Figura 1 Cone of silence: A causa del movimento circolare e dell'inclinazione dell'antenna, un radar è in grado di indagare un'area cilindrica il cui centro è il radar stesso, indicata in verde nell'immagine. È però presente una zona in cui non vengono raccolte informazioni: quest'area ha la forma di un cono rovesciato posto al di sopra dell'antenna, indicata in rosso nella rappresentazione. (radartutorial.eu s.d.)	7
Figura 2 Copertura radar Servizio Meteorologico tedesco: ogni cerchio di colore rosso rappresenta l'area di competenza di ogni antenna radar, con un raggio di 150 km. È importante che vi sia una sovrapposizione delle aree di competenza per non lasciare zone senza informazioni. (radartutorial.eu s.d.)	8
Figura 3 L'"occhio" ben definito e le bande di pioggia (in bianco) dell'uragano Hyacinth, 805km a sud della baia della California, Messico, 09/08/1976. La possibilità di conoscere il movimento della pioggia permette di ricavare la velocità del vento. (NOAA, AP/Wide World Photos s.d.) (Encyclopædia Britannica 2019)	9
Figura 4 Forme differenti di pluviometri standard. Le linee continue indicano l'eventuale presenza di vento, mentre le linee tratteggiate la traiettoria dell'acqua in presenza dello stesso.	13
Figura 5 fattore di correzione k, da applicare per la correzione dell'effetto negativo dato dal vento sulla misura di precipitazione. Vi sono quattro grafici: i due a sinistra riferiti al pluviometro standard "german Hellman", mentre i due a destra al pluviometro a bascula; i grafici superiori si utilizzano per precipitazione liquida, quelli inferiori per precipitazione solida o neve. Inoltre, i simboli vuoti si riferiscono a una pioggia di media durata, quelli pieni per brevi durate, e le linee tratteggiate in maniera differente vengono scelte sulla base della velocità del vento.	14
Figura 6 Pluviometro a bascula: è visibile il pendolo esterno alla bascula (con il magnete) e i contatti che permettono di inviare il segnale quando il reed switch viene attivato	16
Figura 7 Doppia bascula inserita all'interno di un pluviometro, si distingue il meccanismo per convogliare l'acqua all'interno di una delle due vaschette in modo alternato, le quali sono chiuse per evitare un'eccessiva evaporazione. Alla bascula è collegato un magnete, al fondo di un piccolo contrappeso a forma di croce, il quale muovendosi insieme alla bascula attiva il sensore reed switch permettendo la registrazione di un quantitativo fisso di precipitazione proporzionale alla grandezza della bocca di ingresso e della vaschetta.	17
Figura 8 Differenti modelli di stazioni meteo. Da notare in particolare la differente forma dei pluviometri (a sinistra un collettore con bocca da 1000cm ² , a destra invece 500cm ²)	21
Figura 9 Esempio di variazione della risoluzione (Dispense corso GIS 2018, Técnico Lisboa). Una risoluzione spaziale elevata permette un miglior riconoscimento delle forme degli oggetti, ottenendo un maggior dettaglio. La risoluzione spettrale, rappresentata a destra, è legata alla capacità di distinguere valori differenti di una grandezza: in questa rappresentazione si tratta di bande di colore, nel caso oggetto di studio invece, si tratta di considerare intervalli di ampiezza più o meno elevata di lunghezza d'onda delle radiazioni elettromagnetiche.	27
Figura 10 Spettro elettromagnetico. La stima di precipitazione considerata in questo studio è ricavata dalla misura della radiazione con lunghezza d'onda contenuta nelle bande delle microonde e degli infrarossi (Luca Fornaciari Fotografia s.d.)	28
Figura 11 Schema della costellazione GPM, Satelliti principali e bandiere delle nazioni partner (NASA 2015).	30
Figura 12 Rappresentazione del satellite GPM Core Observatory, con indicazione degli strumenti e del campo di azione: il sistema radar agisce tramite "spazzate" con ampiezza 120 o 245 km a terra, mentre il radiometro a microonde con una ampiezza di 885 km. È inoltre indicata la direzione e la velocità di moto e la posizione degli strumenti principali. A destra un ingrandimento del GMI	31
Figura 13 Diagramma dello scanning conico del GMI: Ruotando su se stesso ad un'altezza di 407 km dal suolo, il radiometro in quadra ad ogni rotazione l'area circoscritta dalla linea tratteggiata, composta da numerose misure istantanee circolari incidenti a terra con un angolo di 52,8°. Il radiometro effettua un giro completo 32 volte ogni minuto (NASA 2015).	32
Figura 14 Sopra: rappresentazione del numero di stazioni in gestione alla 3a durante gli anni. Sotto: la disponibilità temporale dei dati della rete pluviometrica e quelli processati da IMERG. Fino a metà 2015 IMERG ha riprocessato dati provenienti dalla missione TRMM, e successivamente quelli raccolti dalla missione GPM	38

Figura 15 Posizioni delle centraline della rete pluviometrica, rappresentazione rispetto all'altitudine	39
Figura 16 Elenco delle variabili e descrizione delle stesse per ogni prodotto in uscita da IMERG	41
Figura 17 Rete Meteorologica in gestione alla 3a S.r.l.	43
Figura 18 Posizione delle stazioni meteorologiche utilizzate disposte sulla stessa griglia di 0,1 x 0,1° in cui vengono elaborate le stime di IMERG. In giallo le caselle contenenti almeno 1 stazione, in blu le celle senza stazioni, non considerate nelle successive analisi	45
Figura 19 Rappresentazione schematica della costruzione della matrice di analisi. Si tratta di sovrapporre le mappe di distribuzione dei dati orari, in modo da poter consultare un particolare momento estraendo un piano orizzontale, oppure una serie storica in una posizione estraendo una colonna.	46
Figura 20 Risultati della tabella di contingenza (Tabella 3): all'aumentare dell'aggregazione temporale sull'asse delle ascisse diminuisce la probabilità di incorrere in periodi senza precipitazione, per cui i correct negatives subiscono una riduzione; contemporaneamente gli hits aumentano, poiché è sufficiente rilevare una singola precipitazione all'interno del periodo di aggregazione per classificare come corretta la stima. Al di sotto invece si trova l'evoluzione degli indici categorici al variare del periodo di aggregazione: il valore ottimale è pari a 0 per il FAR e pari ad 1 per gli altri indici, evidenziando un comportamento positivo della stima.	53
Figura 21 Risultati della tabella di contingenza e degli indici categorici, entrambi elaborati a partire dai dati della Final Run, il prodotto finale e di maggior accuratezza fornito dalla reiterazione dell'algoritmo IMERG e dalla correzione della stima utilizzando dati di una rete pluviometrica, 3 mesi e mezzo dopo l'acquisizione dei dati. Rispetto agli stessi indici elaborati grazie ai dati della Late Run (12 ore dopo l'acquisizione, senza interazione con rete pluviometrica) e visibili in Figura 20, vi è un miglioramento appena apprezzabile.	54
Figura 22 : Boxplot dell'errore quadratico medio, o RMSE, elaborato per differenti aggregazioni temporali per l'anno 2016. Il tratto rosso al centro indica la mediana, mentre i limiti del rettangolo rappresentano il 25° e il 75° percentile della distribuzione, i tratti orizzontali esterni sono i valori massimo e minimo, mentre il + rosso indica i valori considerati "outliers". Considerando periodi di aggregazione maggiori, la differenza tra valori massimi e minimi aumenta, indicando la scarsa accuratezza della stima di IMERG nel rilevare il quantitativo di precipitazione corretto	56
Figura 23 Distribuzione dei valori di RMSE per l'anno 2016, divisi per aggregazione temporale. Ogni colonna rappresenta lo scarto quadratico medio in mm di pioggia per una cella di area 0,1x0,1°, il cui numero è sull'asse delle ascisse. In rosso è visibile la media. All'aumentare dell'aggregazione troviamo scarti maggiori, tuttavia occorre osservare come alcune delle celle con un elevato RMSE nei valori orari siano le stesse che spiccano nell'aggregazione mensile.	56
Figura 24 Andamento RMSE a confronto	57
Figura 25 Distribuzione spaziale dei valori di RMSE per l'anno 2016: i valori aumentano dal blu al rosso. I valori maggiori di scarto si trovano nella zona a nord-ovest, in prossimità dell'arco alpino, e a sud-est dove le stazioni incluse nella cella hanno un'altitudine maggiore. Al contrario, al centro della distribuzione, dove è prevalente un'area pianeggiante, i valori di scarto sono nettamente inferiori.	58
Figura 26 Visuale ingrandita di alcune stazioni poste a nord: le celle di 0,1° di lato in cui sono state incluse durante la spazializzazione presentano RMSE maggiori e differenze più elevate tra stima e misura della centralina pluviometrica, evidenziando la difficoltà del sistema GPM di stimare correttamente la precipitazione in presenza di aree collinari o di rilievi montuosi.	60
Figura 27 Distribuzione dell'altitudine delle stazioni poste nell'area collinare a sud del Piemonte (colore più scuro) rispetto all'intera rete pluviometrica (colore chiaro)	61
Figura 28 Distribuzione spaziale dell'indice di Pearson per l'anno 2016: allargando il periodo di aggregazione l'indice presenta valori maggiori, prossimi all'unità, evidenziando quindi una quasi perfetta corrispondenza tra variazioni della stima del sistema GPM e della misura pluviometrica. Si noti come le zone in cui l'indice risulta essere inferiore sono ancora quelle poste a nord-ovest, in un'area pedemontana, e quelle nella zona collinare a sud-est. Nell'area centrale del Piemonte, in pianura, l'indice presenta valori elevati.	62
Figura 29 Celle di esempio scelte sulla base della posizione geografica, in modo da considerare in dettaglio la serie storica della precipitazione in aree strategiche.	63
Figura 30 Serie storica e scatter, cella 4, valori orari. Sono presenti alcuni picchi in cui le due sorgenti sono discordi. Nello scatter plot la nuvola di punti è distribuita con valori anche distanti dalla linea di riferimento	64

Figura 31 Serie storica e scatter, cella 4, valori giornalieri. I picchi in cui la stima non rispetta la precipitazione osservata sono in numero inferiore: si può infatti osservare una discreta accuratezza della stima, confermata dalla maggior vicinanza della nuvola di punti alla linea di riferimento nello scatter plot. _____ 65

Figura 32 Serie storica e scatter, cella 4, valori per l'aggregazione tri-giornaliera. Vi è una discreta corrispondenza tra rete pluviometrica e stima da satellite: a conferma di ciò il coefficiente R^2 è elevato. _____ 66

Figura 33 Serie storica e scatter, cella 4, valori mensili. L'andamento mensile osservato dalla rete a terra è rispettato dalla stima effettuata da GPM. Tuttavia, il coefficiente R^2 subisce un decremento a causa del minor numero di valori considerati e alle maggiori differenze in millimetri di pioggia dovute all'aggregazione stessa. _____ 67

Figura 34 Serie storica e scatter, cella 1, valori orari. Sono presenti numerosi picchi in cui le due sorgenti sono discordi. Nello scatter plot la nuvola di punti è distribuita con valori anche distanti dalla linea di riferimento, con un R^2 molto basso. _____ 68

Figura 35 Serie storica e scatter, cella 1, valori giornalieri. I picchi in cui la stima non rispetta la precipitazione osservata sono in numero inferiore, tuttavia ve ne sono alcuni (ad esempio a maggio o novembre) in cui la differenza di stima è molto elevata. Il valore del coefficiente di determinazione rimane ridotto. _____ 69

Figura 36 Serie storica e scatter, cella 1, valori per l'aggregazione tri-giornaliera. È possibile notare con chiarezza un periodo tra gennaio e febbraio dove in assenza di osservazione di precipitazione il sistema GPM ne ha stimato la presenza. Sono presenti numerose differenze tra stima e osservazione causando uno scatter plot in cui la nuvola di punti resta distante e molto ampia nei pressi della linea di riferimento. _____ 70

Figura 37 Serie storica e scatter, cella 1, valori mensili. L'andamento annuale della stima nei valori mensili appare molto differente dalla precipitazione osservata. Il valore di R^2 si mantiene molto ridotto indicando una scarsa accuratezza. _____ 71

Figura 38 Variazione coefficiente di determinazione per l'aggregazione giornaliera, cella 4. Il 2016 è l'anno in cui la prestazione è migliore e la stima è maggiormente simile alla realtà osservata. _____ 72

Figura 39 Variazione coefficiente di determinazione, cella 1. In tutti gli anni il valore di R^2 è tale da considerare negativamente la stima da parte di IMERG. Il 2016 è l'anno peggiore, ma il coefficiente è ridotto in tutti gli anni considerati. _____ 72

Figura 40 Distribuzione spaziale del coefficiente di determinazione per anno 2015. All'aumentare dell'aggregazione notiamo un miglioramento dell'indice. Le celle a nord della regione presentano i valori peggiori, mentre nell'area pianeggiante a sud-ovest si trova un'area in cui R^2 è generalmente al di sopra di 0,6. _____ 74

Figura 41 Risultati della tabella di contingenza per l'anno 2013: all'aumentare della finestra temporale sull'asse delle ascisse diminuisce la probabilità di incorrere in periodi senza precipitazione, per cui i correct negatives subiscono una riduzione; contemporaneamente gli hits aumentano, poiché è sufficiente rilevare una singola precipitazione all'interno del periodo di aggregazione per classificare come corretta la stima. Al di sotto invece si trova l'evoluzione degli indici categorici al variare del periodo di aggregazione: il valore ottimale è pari a 0 per il FAR e pari ad 1 per gli altri indici, evidenziando un comportamento positivo della stima. _____ 83

Figura 42 Risultati della tabella di contingenza 2014: l'evoluzione degli indici categorici e delle categorie prestazionali al variare del periodo di aggregazione temporale rispecchia il comportamento dei medesimi indici nell'anno precedente. _____ 84

Figura 43 Risultati della tabella di contingenza 2015: l'evoluzione degli indici categorici e delle categorie prestazionali al variare del periodo di aggregazione temporale rispecchia il comportamento dei medesimi indici nell'anno precedente. Tuttavia, è possibile notare un minor successo negli indici categorici, evidente nei valori percentuali inferiori nell'aggregazione mensile. _____ 84

Figura 44 Risultati della tabella di contingenza 2017: l'evoluzione degli indici categorici e delle categorie prestazionali al variare del periodo di aggregazione, a differenza degli anni precedenti, presenta una maggior percentuale di false alarms. Tuttavia, in generale il comportamento è tale da poter essere accostato a quello degli altri anni considerati nell'analisi. _____ 85

Figura 45 Distribuzione spaziale dei valori di RMSE 2013: i valori aumentano dal blu al rosso. I valori maggiori di scarto si trovano nelle zone a nord ed a ovest, in prossimità dell'arco alpino, dove le stazioni incluse nella cella hanno un'altitudine maggiore. Vi è inoltre una piccola area a sud in cui

la topografia è collinare e lo scarto risulta maggiore. Al contrario, al centro della distribuzione, dove prevale la pianura, i valori di scarto sono nettamente inferiori. _____ 85

Figura 46 Distribuzione spaziale dei valori di RMSE 2014: si può notare la presenza di valori maggiori nelle aree limitrofe a nord, ovest e sud-est, in cui la vicinanza all'arco alpino e la topografia di natura collinare produce errori di stima da parte di IMERG. _____ 86

Figura 47 Distribuzione spaziale dei valori di RMSE 2015: rispetto all'anno precedente i valori di scarto si presentano minori, con ampie aree nelle quali RMSE si mantiene inferiore a 10mm nelle aggregazioni giornaliere e tri-giornaliere. _____ 86

Figura 48 Distribuzione spaziale dei valori di RMSE per l'anno 2017: è possibile osservare valori di scarto inferiori agli anni precedenti, in particolare rispetto al periodo 2013-2015. Un possibile spiegazione può essere attribuita al definitivo utilizzo di satelliti della missione GPM ed all'abbandono di quelli di competenza di TRMM. Anche in questo caso, tuttavia, la stima di pioggia presenta uno scarto elevato nella zona settentrionale della distribuzione, mentre al centro si nota un solo valore anomalo. _____ 87

Figura 49 Distribuzione spaziale dell'indice di Pearson 2013: aumentando il periodo di aggregazione, l'indice presenta valori maggiori, prossimi all'unità in taluni casi, evidenziando quindi una quasi perfetta corrispondenza tra variazioni della stima del sistema GPM e della misura pluviometrica. Si noti come le zone in cui l'indice risulta essere inferiore sono ancora quelle poste a nord-ovest, in un'area pedemontana, e quelle nella zona collinare a sud-ovest. L'indice presenta valori maggiori nelle aggregazioni di 3 giorni rispetto a quelle mensili. _____ 87

Figura 50 Distribuzione spaziale dell'indice di Pearson 2013: confrontando la distribuzione con quella dell'anno precedente si sottolinea un generale miglioramento, con valori nell'area collinare a sud-est prossimi all'unità. Questo significa che vi è buona correlazione tra le serie temporali di osservazione e stima. _____ 88

Figura 51 Distribuzione spaziale dell'indice di Pearson 2015: si noti il miglioramento dell'indice fino all'aggregazione tri-giornaliera, presentando poi valori inferiori nell'aggregazione mensile. Tuttavia, i valori dell'indice di Pearson sono elevati in tutta l'area meridionale del Piemonte, presentando ancora un'eccezione nelle celle a nord-ovest prossime ai rilievi alpini. _____ 88

Figura 52 Distribuzione spaziale dell'indice di Pearson 2017: si osservi il migliorare dei valori dell'indice all'aumentare dell'aggregazione: vaste aree nelle aggregazioni dai 3 giorni in su si presentano superiori a 0,8. In questo anno, le aree a prestazioni minori sono quelle site a sud ovest, anch'esse prossime all'arco alpino. _____ 89

Figura 53 Distribuzione spaziale del coefficiente di determinazione 2013: vi è un aumento dei valori di R^2 con l'aumento dell'aggregazione temporale nei giornalieri e tri giornalieri, per poi subire una ricaduta nell'aggregazione mensile. In generale il coefficiente si mantiene inferiore a 0,6 in tutta la distribuzione. _____ 89

Figura 54 Distribuzione spaziale del coefficiente di determinazione 2014: si noti il costante miglioramento dell'indice in contemporanea all'aumento del periodo di aggregazione. Nell'area collinare a sud est, generalmente a prestazioni negative negli altri indici, il coefficiente R^2 presenta valori positivi prossimi all'unità. _____ 90

Figura 55 Distribuzione spaziale del coefficiente di determinazione 2016: si notino valori positivi e prossimi a 1 in tutta l'area meridionale della distribuzione. Tuttavia, nell'aggregazione mensile vi è una leggera riduzione di R^2 . Le aree in cui il coefficiente si presenta con valori ridotti si trovano a nord ovest ed a sud est, rispettivamente in prossimità dell'arco alpino e in zona collinare. _____ 90

Figura 56 Distribuzione spaziale del coefficiente di determinazione 2017: rispetto agli anni precedenti i valori di R^2 appaiono nettamente inferiori in tutte le aggregazioni, pur presentando numerose celle in cui il coefficiente è superiore a 0,6. _____ 91

1. Introduzione

L'acqua è l'elemento che governa l'ambiente in cui viviamo: dall'alimentazione alla produzione industriale, l'acqua è in continuo contatto con l'uomo, ne costituisce il fondamento della vita. Con l'introduzione dei *Sustainable Development Goals* da parte delle Nazioni Unite la comunità globale ha messo in evidenza l'importanza di una corretta gestione delle risorse idriche per combattere il cambiamento climatico e raggiungere la sostenibilità ambientale, dedicando uno dei 17 targets alla sola acqua. (United Nations 2015)

Per poter gestire nel modo appropriato una risorsa così importante, è innanzitutto necessario disporre di quante più informazioni possibili, ovvero avvalersi di tecnologie efficaci e di investimenti oculati per promuovere un utilizzo consapevole. Kofi Annan, ex segretario generale dell'ONU, nel discorso del 2018 alla World Meteorological Organisation, affermò che *“l'acqua è una risorsa critica per la salute umana, la sicurezza degli alimenti, l'ambiente, la stabilità e la prosperità; abbiamo di conseguenza il dovere di porre l'acqua al centro di tutti gli sforzi volti al raggiungimento della pace e dello sviluppo”*. (Annan 2018)

Il primo passo è dunque la conoscenza dei fenomeni tramite il monitoraggio ambientale.

La precipitazione ricopre un ruolo di notevole importanza: si tratta di una fase cruciale del processo di trasporto dell'acqua, nella quale il vapore disperso in atmosfera si condensa in nubi e vi può essere una conseguente precipitazione a terra in forma liquida o solida. Il ciclo idrico è noto nella sua interezza, ma i movimenti delle grosse masse d'acqua sono ancora oggi oggetto di discussione e verifica. Per valutare i volumi in gioco, acquista dunque particolare importanza la possibilità di misurare la precipitazione con precisione. Determinare la quantità di acqua caduta al suolo in una specifica zona è utile tanto in ambito ingegneristico a scopo di progettazione, ad esempio, di opere idrauliche o di protezione dal dissesto idrogeologico, quanto in altre attività quali l'agricoltura, la protezione civile, la tutela del patrimonio naturale. Nel 1970 il Club di Roma pubblicò *“I limiti dello sviluppo”* (Meadows, et al. 1972), un rapporto su come il pianeta Terra non sia in grado di sopportare il ritmo di crescita economica e demografica dell'uomo, ponendo l'attenzione per la prima volta sul cambiamento climatico, ma mai come oggi esso è diventato un argomento di rilevanza globale. Monitorare le condizioni ambientali permette di conoscere lo stato attuale, e prevedere con più precisione il futuro, ambendo a raggiungere la sostenibilità tramite una migliore gestione delle risorse.

Vi è un numero incredibile di data provider grazie ai quali ottenere parametri climatici utili, ognuno dei quali possiede sue peculiarità: dal tipo di tecnologia utilizzata, alla diffusione e accessibilità dei dati, alla rapidità con la quale questi dati sono disponibili. Infatti, alcuni enti principalmente pubblici, forniscono libero

accesso ai loro dati, talvolta in real-time, altre volte con tempi più dilazionati, o ancora forniscono i risultati delle *reanalysis*, ovvero rielaborazioni dei dati raccolti in passato, resi così più precisi e affidabili.

L'intento di questa tesi è quello di confrontare due differenti tipologie di misura della precipitazione diametralmente opposti: quella effettuata tramite l'utilizzo di un pluviometro, o meglio, una rete di pluviometri disposti su una rete locale, e quella stimata dall'analisi di numerosi parametri raccolti da una costellazione di satelliti. La scelta di queste due fonti non è casuale: in parte è dettata dalla disponibilità di un'azienda di Torino attiva nella fornitura di servizi meteorologici e di supporto all'agricoltura, la 3a S.r.l., a fornire i dati raccolti dalla rete pluviometrica che ha in gestione, in parte deriva dagli aspetti che queste tecniche hanno in comune. Entrambe le sorgenti considerano dati orari e sub-giornalieri, utili in numerose applicazioni per processi con un'evoluzione rapida, come può essere la precipitazione durante un temporale; questi dati sono consultabili in remoto, grazie ad un database che si aggiorna automaticamente e sfrutta la tecnologia cloud e la rete internet mobile; sono entrambi sistemi attivi da anni per cui si dispone di una serie storica considerevole. I dati sono raccolti in continuo ed è possibile utilizzarli in tempo quasi-reale, con un intervallo minimo di pochi minuti per la rete a terra e di 4 ore per il satellite, suggerendo quindi la possibilità di sfruttare entrambe le osservazioni in maniera integrata. Dall'altro lato le due sorgenti si differenziano considerevolmente: una rete pluviometrica rappresenta una serie di valori puntuali, distribuiti eterogeneamente sul territorio, mentre la posizione sopraelevata del satellite fa sì che venga considerata in un solo istante un'area molto più ampia e "continua", senza gli ostacoli che vi possono essere a terra; la misura del pluviometro è dunque materiale, mentre riguardo al satellite si tratta di una stima basata sull'interazione dell'acqua meteorica con le radiazioni elettromagnetiche provenienti dalla superficie terrestre. Tale stima rappresenta il prodotto finale di un algoritmo sviluppato all'interno della *Global Precipitation Measurement* (GPM), una recente missione di NASA e JAXA, rispettivamente gli enti spaziali americano e giapponese. Per meglio comprendere l'offerta in fatto di strumenti per la misura della precipitazione è stata fatta un'analisi dello stato dell'arte delle principali tecnologie per la misura, evidenziandone gli aspetti positivi e negativi, per poi focalizzare lo studio sui metodi considerati maggiormente importanti.

Inoltre, si è scelto di considerare il territorio della Regione Piemonte come area di studio: si tratta infatti di un'area particolare, composta da rilievi montuosi elevati che circondano un'area pianeggiante che si allarga ad est nella pianura padana, dando luogo al clima temperato tipico dell'Italia settentrionale. Diversi studi hanno analizzato il comportamento del GPM in aree con orografia complessa (Caracciolo 2018) (Xu 2016), portando a risultati controversi e suggerendo che l'accuratezza dipenda da molteplici fattori. Stabilire quindi quale sia la reale portata di un sistema in remoto, il suo comportamento in presenza di variazioni

topografiche considerevoli in un'area ridotta, e la possibilità che esso possa sostituire una rete pluviometrica a terra con un'accuratezza pari o confrontabile sono i quesiti principali a cui questo studio ambisce a rispondere.

2. Stato dell'arte per le tecnologie di misurazione della precipitazione

L'importanza di disporre di una misura accurata della precipitazione è fondamentale in numerose applicazioni scientifiche: la conoscenza è infatti il primo passo per una corretta gestione delle risorse idriche, ovvero la capacità di creare progetti futuri in cui avvalersi della capacità della precipitazione di ricaricare corpi idrici superficiali e sotterranei. Lo studio del clima è necessario, ad esempio dal punto di vista ingegneristico, per la prevenzione del dissesto idrogeologico, la gestione degli approvvigionamenti idrici, la progettazione di opere di difesa spondale. Anche nelle applicazioni agricole la conoscenza della quantità di acqua caduta al suolo, delle altre idrometeorie e dei parametri climatici riveste una notevole importanza per applicare trattamenti e azioni di irrigazione mirate ed ottenere una produzione più efficace.

Per poter misurare con precisione la quantità di pioggia che cade al suolo è necessario considerare che la precipitazione può avere natura liquida o solida: pioggia, grandine o neve. Tuttavia, siccome si è scelto di operare questo studio concentrandosi su una zona temperata, il territorio piemontese, e su un'area a vocazione agricola, la maggior parte delle precipitazioni prese in considerazione sono da intendersi come liquide.

Verrà presentata di seguito una breve rassegna della storia e dell'evoluzione dei principali strumenti di cui l'uomo si è avvalso nel tempo e si avvale tuttora: dai primi strumenti, somiglianti maggiormente a semplici contenitori aperti, i quali potevano effettuare misure con un intervallo temporale variabile in base alle necessità, all'evoluzione della tecnologia che oggi ci spinge ad effettuare misure in remoto, rendendo più corretto riferirsi ad essa come *stima* o *calcolo* della pioggia.

Tra gli strumenti più validi ed utilizzati in campo vi è il pluviometro, mentre il sistema satellitare e altri dispositivi in orbita sono oggi una valida alternativa per chi necessita di una misura con maggior estensione spaziale. Il pluviometro, concettualmente semplice e molto diffuso, è lo strumento più sicuro ed efficiente: viene infatti utilizzato nelle operazioni di validazione di altre tecnologie. Dall'altro lato i satelliti possono essere di differente natura, raccogliere informazioni variegata e sono in grado di utilizzare un set di strumenti integrati su un'unica piattaforma per indagare lo stesso fenomeno sull'intera superficie terrestre.

Esistono tuttavia alcune vie intermedie: il radar è un sistema la cui misura a medio raggio (tra i 150 e i 250km) è oggi consolidata, per questo ampiamente utilizzata in numerosi servizi meteorologici. In aggiunta a questi metodi ormai consolidati, ve ne sono altri maggiormente sperimentali e numerose tecnologie ibride di interpolazione di misure pluviometriche provenienti da differenti fonti: ad esempio all'interno del progetto europeo Copernicus troviamo ERA5, un dataset in continuo aggiornamento derivato dall'analisi e rianalisi di dati ambientali

raccolti sia tramite satellite che in-situ (ECMWF 2020). Un altro esempio degno di nota riguarda un recente progetto italiano che si avvale di un algoritmo (SM2RAIN, Soil Moisture To Rain) in grado di generare una stima di precipitazione a partire dall'umidità del suolo misurata da satellite (Chiaravalloti 2018).

La qualità di acquisizione di un pluviometro è limitata dalla puntualità dell'acquisizione. Nel campo della meteorologia è dimostrato come i valori di una stessa grandezza possano variare con velocità sorprendente, ma non solo: all'interno dell'atmosfera si possono distinguere numerosi strati, ed ognuno di essi ha caratteristiche e, spesso, composizioni differenti. La variabilità delle condizioni atmosferiche è, soprattutto per il nostro interesse, orizzontale: dislivelli, vegetazione, installazioni antropiche e numerosi altri elementi possono modificare localmente le condizioni di temperatura e umidità. In quest'ottica risulta utile avvalersi in alcune applicazioni degli strumenti con un respiro più ampio, provvisti della capacità di cogliere e misurare i fenomeni di interesse in un'area di maggior estensione. Per ottenere una visione d'insieme più ampia delle variazioni di una grandezza, è tuttavia imprescindibile adottare un punto di osservazione sopraelevato e a maggior distanza dal luogo di osservazione. Rendendosi necessario questo distacco, è possibile che la misura sia soggetta a una perdita di accuratezza: infatti si considerano come *certi* i valori di precipitazione registrati dai pluviometri, mentre le misure effettuate da radar o satellite necessitano di diverse analisi in post-processing prima di poter essere sfruttate con valenza scientifica.

2.1. Radar a terra

Un radar, acronimo di *R*Adio *D*etecting *A*nd *R*anging (radiatorilevamento e misurazione a distanza), è un sistema indiretto di misura delle precipitazioni, liquide o solide che siano. Più in generale, i radar sono in grado di sfruttare le onde elettromagnetiche per individuare un oggetto posto a distanza: un segnale elettromagnetico di frequenza nota è inviato in una specifica direzione, interferisce con gli oggetti con cui viene a contatto e parte di esso viene riflesso tornando all'origine. La radiazione di ritorno è rilevata da un'antenna e interpretata: grazie ad alcuni elementi noti (frequenza dell'onda, velocità di spostamento nel mezzo, tempo di ritorno) è possibile risalire alla dimensione e alla posizione relativa dell'oggetto di indagine. Infatti, la radiazione elettromagnetica subisce diversi fenomeni interagendo con la materia, tra cui una dispersione quando colpisce un oggetto di dimensione ad una scala simile alla propria lunghezza d'onda (*backscattering*, o retrodiffusione), permettendo al segnale di tornare indietro nella stessa direzione e di essere captato dall'antenna ricevente. In virtù di queste qualità, oltre a ricavare la presenza, la posizione e la velocità degli oggetti, un radar può, in alcuni casi, ricavarne la dimensione e la forma. (Goudenhoofdt 2009)

Ai fini di questa tesi, è opportuno indagare più nello specifico il funzionamento di un radar meteorologico. I radar appartenenti a questa categoria, servendosi di un'antenna in rotazione, inviano in molteplici direzioni un segnale elettromagnetico di brevissima durata (di solito si tratta di un impulso di 1 millisecondo) la cui lunghezza d'onda si trova nella banda propria delle microonde. Successivamente, per un ulteriore tempo di 1ms l'antenna rimane in funzione ricevente e attende il ritorno del segnale. Le lunghezze d'onda dell'impulso sono scelte in modo da interagire correttamente con le eventuali gocce di acqua o grandine presenti e in modo che possa verificarsi il *backscattering*. La distanza, infine, è misurabile sapendo che l'impulso si muove alla stessa velocità della luce, ovvero $3 \cdot 10^8$ m/s, sfruttando le leggi fondamentali del moto.

Il radar fisicamente si compone di tutti gli elementi necessari alla creazione dell'impulso (principalmente un *klystron*, ovvero uno strumento per generare un impulso di natura elettromagnetica ed in grado di convogliarlo in una direzione e di un clock di riferimento), di un'antenna le cui dimensioni variano in base alla potenza del segnale, e infine di una cupola di protezione: infatti per indagare tutto l'orizzonte, l'antenna ruota su se stessa e invia il medesimo impulso a raggiera attorno a sé, perciò per la sicurezza degli operatori è coperto. Inoltre, solitamente è possibile inclinare l'antenna per sondare con più precisione strati superiori dell'atmosfera: questo movimento evita in parte il crearsi di una zona conica eccessivamente ampia posta al di sopra dello strumento su cui non ci sarebbe misurazione (*cone of silence*, Figura 1), a causa dell'espansione su un fronte semicircolare dell'impulso (radartutorial.eu s.d.).

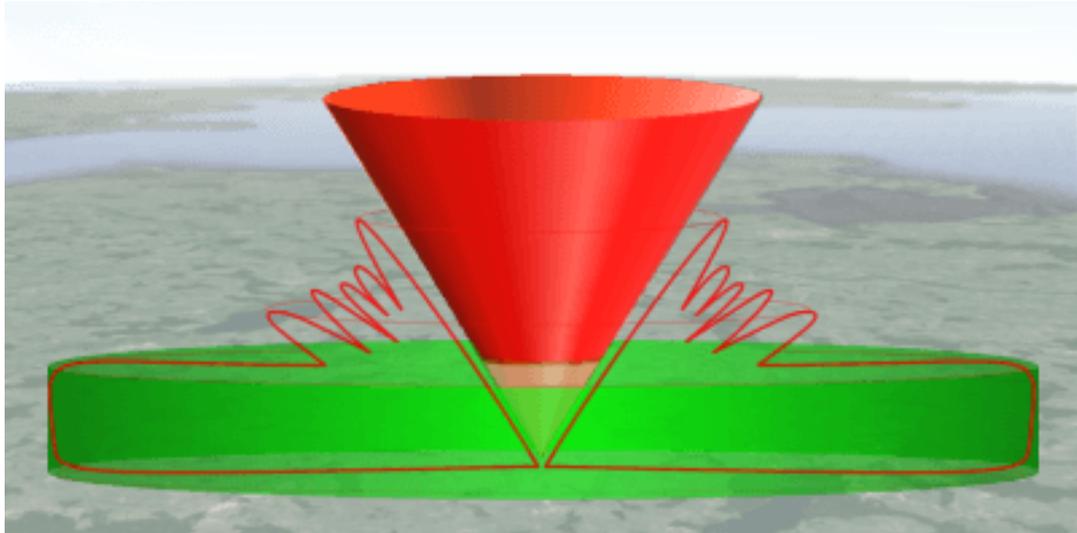


Figura 1 Cone of silence: A causa del movimento circolare e dell'inclinazione dell'antenna, un radar e in grado di indagare un'area cilindrica il cui centro è il radar stesso, indicata in verde nell'immagine. È però presente una zona in cui non vengono raccolte informazioni: quest'area ha la forma di un cono rovesciato posto al di sopra dell'antenna, indicata in rosso nella rappresentazione. (radartutorial.eu s.d.)

La capacità di un radar non è determinata esclusivamente dalla strumentazione, bensì anche dalla distanza a cui è desiderio indagare. Infatti, l'impulso inviato si diffonde secondo una superficie sferica, perdendo proporzionalmente intensità man mano che occupa un volume d'aria maggiore. Inoltre, l'insorgenza di interferenze con altri segnali e la presenza di elementi orografici ed installazioni umane possono condizionare il segnale ricevuto in risposta all'impulso: il radar ha quindi un'elevata accuratezza a corto raggio ma perde rapidamente di dettaglio, per questo nelle reti meteorologiche è importante avere una sovrapposizione delle aree di misura. Ad esempio, il servizio meteorologico nazionale tedesco (Deutscher wetterdienst s.d.) copre il territorio di sua competenza con 17 strumenti, ognuno operante su un raggio di 150km.

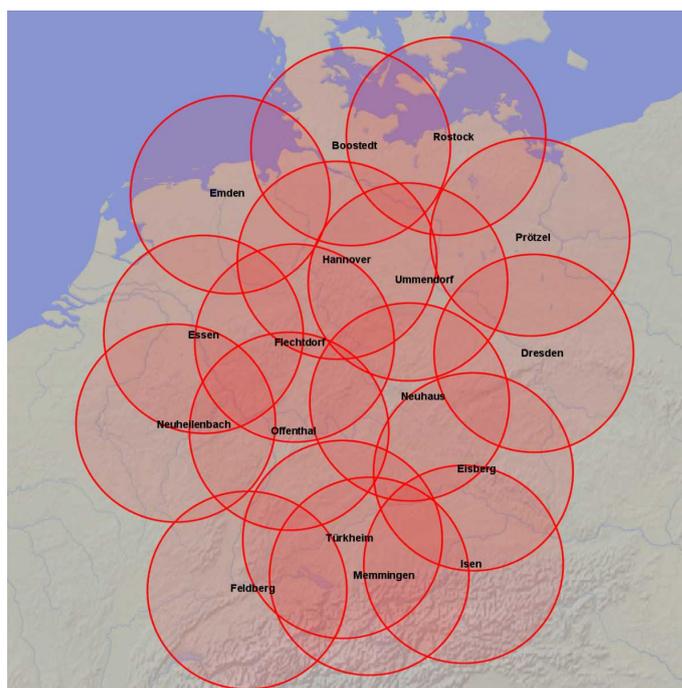


Figura 2 Copertura radar Servizio Meteorologico tedesco: ogni cerchio di colore rosso rappresenta l'area di competenza di ogni antenna radar, con un raggio di 150 km. È importante che vi sia una sovrapposizione delle aree di competenza per non lasciare zone senza informazioni. (radartutorial.eu s.d.)

Esistono due tipi principali di radar impiegati in ambito meteorologico: polarimetrici e Doppler.

Un radar polarimetrico è in grado di misurare tutte le caratteristiche che dipendono dalla polarizzazione del segnale: può quindi inviare e ricevere impulsi con orientazione sia verticale che orizzontale, oppure ricevere con un'orientazione differente da quella inviata. In questo modo è possibile misurare e prevedere la forma e la grandezza dell'obiettivo su più dimensioni. Il principale vantaggio è di poter escludere elementi le cui dimensioni non sono coerenti con una precipitazione, distinguere tra pioggia "pesante" e quella "leggera", o in altre parole, tra pioggia, grandine o neve.

I radar del secondo tipo indicato sono invece in grado di misurare lo slittamento di frequenza (meglio noto come *Doppler frequency shift*) oltre alla magnitudo del segnale riflesso dalla precipitazione. L'importanza dello shift è legata alla velocità radiale delle gocce mosse dai venti, che si allontanano o avvicinano al radar doppler. Considerando che tornado, cicloni, uragani e altri fenomeni simili seguono un moto rotatorio di vario genere, risulta molto importante la possibilità di misurare la velocità delle gocce (quindi indirettamente quella del vento) e la struttura del vortice tramite il moto della precipitazione.

Come detto in precedenza, l'invio di impulsi in direzioni e quote diverse permette di creare una visualizzazione 3D della precipitazione, quindi di misurare anche gli angoli con cui il vento spirava e, conseguentemente, il raggio e il tipo di moto radiale di fenomeni intensi.



Figura 3 L'occhio ben definito e le bande di pioggia (in bianco) dell'uragano Hyacinth, 805km a sud della baia della California, Messico, 09/08/1976. La possibilità di conoscere il movimento della pioggia permette di ricavare la velocità del vento. (NOAA, AP/Wide World Photos s.d.) (Encyclopædia Britannica 2019)

Infine, prima di poter sfruttare i radar appieno, è necessario effettuare una calibrazione per un periodo sufficientemente lungo, seguita da una validazione. Nonostante ciò, spesso i risultati del radar meteo sono corretti e modificati avvalendosi di una rete di pluviometri. Molto sovente i termini calibrazione, validazione e correzione sono confusi tra loro, perciò per chiarezza verranno enunciate le loro definizioni sintetiche:

- *Calibrazione*: processo il cui obiettivo finale è assicurare la stabilità e riproduzione delle misure radar a lungo termine, come spiegato chiaramente da U. Gjertsen, M. Štek, e D. B. Michelson (Gjertsen 2004). La procedura è elettronica e viene spesso effettuata dal sistema radar stesso con cadenza periodica.
- *Validazione*: è un vero e proprio controllo successivo alla calibrazione. Si misura uno stesso fenomeno con lo strumento appena calibrato e con quello che riteniamo effettui la misura esatta, il quale può essere a sua volta uno strumento calibrato e già validato. L'obiettivo è di verificare la qualità della calibrazione e, ove necessario, effettuarla nuovamente.
- *Correzione*: in particolare la correzione dei radar tramite rete pluviometrica, è una procedura statistica atta a raggiungere una stima di precipitazione più accurata comparando ad una rete di pluviometri. Il risultato finale rappresenta dunque un valore ricavato dalla misura radiometrica e modificato in modo da essere eguale a quello pluviometrico. Ci sono numerosi studi a riguardo e molteplici metodologie tramite le quali è

possibile operare simili correzioni, sulle quali non ci si soffermerà in questo studio (Gjertsen 2004) (Mahadi Hasan 2015).

2.2. Pluviometri

Il pluviometro è lo strumento al momento più semplice ed efficace per la misura della quantità di pioggia che cade al suolo. Ideato e utilizzato fin dal medioevo, fino a 20 anni fa il pluviometro era costituito un semplice imbuto collegato ad un cilindro graduato che richiedeva quotidianamente la lettura e lo svuotamento, senza permettere di indagare l'evoluzione della precipitazione durante la giornata, a meno di svuotarlo regolarmente.

Grazie all'introduzione di sistemi differenti, è diventato possibile conoscere la quantità di acqua precipitata in un determinato periodo di tempo inferiore al giorno (o al periodo di svuotamento del contenitore). Negli ultimi 30 anni differenti forme del meccanismo di raccolta, sistemi in grado di svuotarsi automaticamente e tecnologie per registrare associate allo strumento di misura hanno reso realtà lo studio dell'intensità di precipitazione (indicato spesso con il termine inglese *rain rate*).

Se associato ad un sistema di registrazione il pluviometro cambia locuzione in *pluviografo*: i primi e più rudimentali utilizzavano come strumento di scrittura un pennino e della carta avvolta attorno ad un rullo, perciò ad ogni movimento conseguente all'immissione di acqua nel sistema, tipicamente con scatti di 0,2 mm, veniva segnato un trattino sulla carta, mentre in contemporanea il rullo ruotava allo scorrere del tempo. In particolare, è noto che l'unità di misura universalmente adottata è il millimetro: 1mm di pioggia corrisponde a 1litro/m², permettendo di risalire ai volumi molto facilmente. Oggi il termine pluviometro e pluviografo sono interscambiabili, poiché la quasi totalità possiede un sistema di registrazione della precipitazione.

Con il progredire dell'elettronica e con l'introduzione della rete internet mobile, è oggi possibile misurare l'accumulo di pioggia, la sua distribuzione temporale e quindi il totale con diversi tempi di aggregazione. In ambito agricolo in genere lo strumento si aggiorna ogni minuto e può produrre valori di aggregazione dai 5 minuti in su, così che sia possibile misurare non solo la quantità ma anche il tasso di precipitazione, utile per esempio per valutare la capacità di assorbimento del terreno.

Inizialmente, le misure erano registrate su una scheda di memoria, come sostituto digitale del rullo di carta e del pennino. Oggi, tuttavia, la tendenza generale è di inviare i dati ad un database centrale, ove possibile, tramite la rete telefonica.

2.2.1. Caratteristiche tecniche

L'Organizzazione Meteorologica Mondiale o WMO (World Meteorological Organisation 2017) ha stabilito degli standard per la misura della precipitazione: l'ultima versione della "*WMO guide to meteorological instruments and methods of observation*" risale al 2014, aggiornata e corretta nel 2017. In questo modo, le misure provenienti da qualsiasi parte del globo possono essere certificate e utilizzate con uno stesso livello minimo di accuratezza. Al capitolo 6 della parte I della guida sono contenute una serie di direttive affinché una precipitazione possa essere misurata in maniera corretta. Di seguito sono sintetizzate le condizioni basilari da rispettare.

- La precipitazione è definita come "il prodotto liquido o solido della condensazione del vapore che precipita da nuvole o si deposita dall'aria sul terreno". In base a questa definizione è considerata precipitazione: pioggia, neve, grandine, rugiada, brina, nebbia depositata. L'unità di misura è la profondità dell'acqua distribuita su una superficie nota (millimetri su metro quadro, ovvero mm/m²). La risoluzione minima accettabile è 0,2 mm, oppure 0,1 "se fattibile". La misura dell'intensità o tasso di precipitazione è in mm/h (mm*h⁻¹).
- La misura è molto sensibile all'esposizione. Vanno infatti tenuti in considerazione molti aspetti che concorrono a produrre valori esatti. Nonostante la misura effettuata con pluviometro sia la risorsa primaria per l'analisi areale delle precipitazioni, si tratta di una misura puntuale, rappresentativa di un'area proporzionale al tempo di osservazione e alla topografia locale. In linea di principio radar e satelliti sono un'alternativa per definire la distribuzione spaziale delle piogge, tuttavia sarebbe auspicabile un'integrazione delle tre risorse per una misura affidabile e usufruibile con solida sicurezza. I pluviometri, nonostante questa ideale interazione, hanno il dovere di produrre una misura rappresentativa della quantità di acqua caduta al suolo nella zona, non solo nel punto di installazione: si rende necessario un posizionamento studiato e ottimale, pena un errore sistematico come evidenziato da Sevruk and Zahlavova (Sevruk e Zahlavova 1994). Gli elementi che possono portare complicazioni maggiori sono indubbiamente la presenza di vento ed oggetti che possono ostacolare la raccolta di acqua nel collettore del pluviometro, quali vegetazione troppo alta od edifici. La WMO consiglia di posizionare lo strumento ad una distanza pari ad almeno il doppio dell'altezza della bocca rispetto a qualsiasi ostacolo, in posizione piana e non in cima ad edifici, quindi al suolo o a poca distanza da esso, in modo da evitare l'in-spashing, ovvero l'ingresso di gocce o piccole quantità d'acqua nello strumento in seguito al rimbalzo su altre superfici. Per questo motivo l'eventuale

vegetazione attorno al pluviometro deve essere potata in modo da rimanere al di sotto del collettore di ingresso, oppure bisogna ricorrere all'utilizzo di paravento che la isolino. Oltre a ciò è importante che l'ingresso del pluviometro abbia bordi rialzati. In Figura 4 si riportano le forme dei pluviometri standard accettati dalla WMO: osservando le linee orizzontali rappresentanti il vento occorre prestare attenzione alla deformazione dei flussi subito nei pressi dell'apertura del sistema di raccolta. Ad un diametro minore si associano deformazioni inferiori e, auspicabilmente, anche gli errori collegati alla presenza del vento saranno più piccole. Quest'ultimo è infatti la principale causa di sottostima della precipitazione.

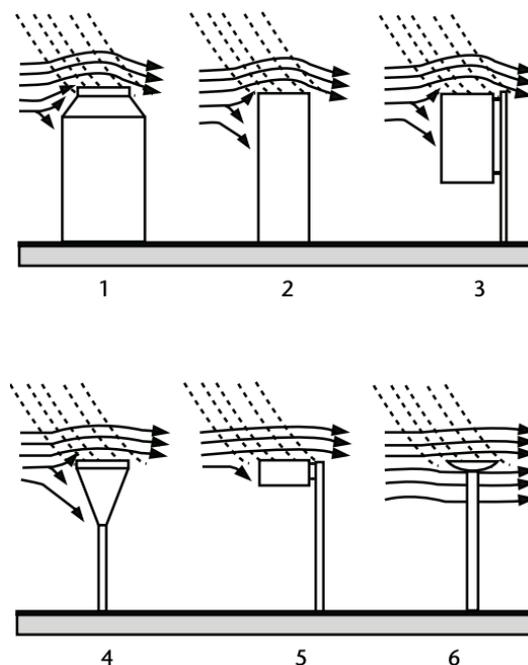


Figura 4 Forme differenti di pluviometri standard. Le linee continue indicano l'eventuale presenza di vento, mentre le linee tratteggiate la traiettoria dell'acqua in presenza dello stesso.

- L'area della bocca deve essere costante, ovvero non subire variazioni dovute alla temperatura superiori allo 0,5%. Deve impedire l'entrata e (l'uscita) di gocce tramite rimbalzo involontario: questo si può ottenere adottando un bordo rialzato, un'inclinazione della parete interna della bocca superiore al 45% e grazie alla scelta del materiale più adatto in modo che l'acqua possa scivolare verso lo strumento di misura evitando l'evaporazione. Ovviamente è necessario calibrare lo strumento eseguendo prove di immissione di volumi noti di acqua.

In generale, sarebbe auspicabile evitare di dover correggere gli errori a posteriori rispetto alla misura, poiché sono necessari numerosi dati supplementari quali: la velocità del vento, la dimensione delle gocce, il tipo di precipitazione, l'intensità di precipitazione, temperatura e umidità dell'aria, caratteristiche aggiuntive della

posizione del pluviometro. Nel 1982 la WMO propose una prima formula di correzione, poi modificata nel 1990 da Legates and Willmott (Willmott e Legates 1990):

$$P_k = k_r P_{cr} + k_s P_{cs}$$

dove

$$P_{cr} = P_{gr} + \Delta P_{1r} + \Delta P_{2r} + \Delta P_{3r} + \Delta P_{4r}$$

$$P_{cs} = P_{gs} + \Delta P_{1s} + \Delta P_{2s} + \Delta P_{3s} + \Delta P_{4s}$$

e

- P_k è la pioggia corretta
- r ed s sono “(liquid) rain” e “(solid) snow”
- k è il fattore di correzione dell'effetto di deformazione del vento (cfr. Figura 5)
- P_c è la precipitazione raccolta dalla bocca/collettore
- P_g è la precipitazione misurata
- ΔP_1 è la perdita di precipitazione nelle pareti interne del collettore
- ΔP_2 è la perdita di precipitazione nella bascula dopo lo svuotamento
- ΔP_3 è la perdita di precipitazione per evaporazione
- ΔP_4 è la perdita di precipitazione per errori meccanici sistematici

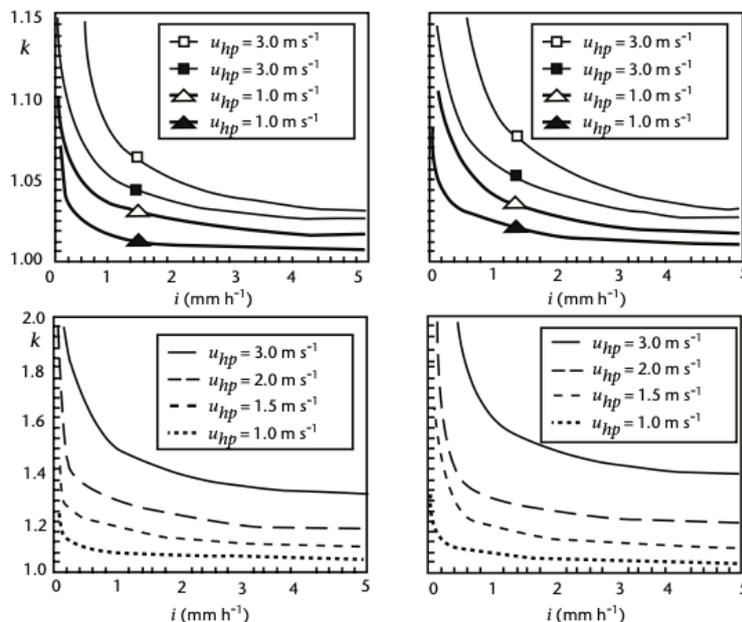


Figura 5 fattore di correzione k , da applicare per la correzione dell'effetto negativo dato dal vento sulla misura di precipitazione. Vi sono quattro grafici: i due a sinistra riferiti al pluviometro standard “german Hellman”, mentre i due a destra al pluviometro a bascula; i grafici superiori si utilizzano per precipitazione liquida, quelli inferiori per precipitazione solida o neve. Inoltre, i simboli vuoti si riferiscono a una pioggia di media durata, quelli pieni per brevi durate, e le linee tratteggiate in maniera differente vengono scelte sulla base della velocità del vento.

Il fattore di correzione k indica il rapporto tra precipitazione misurata e corretta, e la sua versione più recente è stato ricavata sperimentalmente nei laboratori della WMO nel 2009. Esso dipende dalla velocità del vento (u_{hp}), dall'intensità di precipitazione i , dal tipo di pluviometro e dalla modalità di precipitazione (Sevruk e Nespor 1999). In Figura 5 si trovano quattro grafici: i due a sinistra si riferiscono al pluviometro standard "German Hellmann" (del tipo a cilindro graduato) e i due a destra al pluviometro a bascula; mentre i due superiori si applicano nel caso di precipitazione liquida (pioggia), i due sotto servono per la precipitazione solida (neve). I simboli vuoti rappresentano una pioggia di media durata, quelli pieni invece si applicano per brevi durate (*showers*). Occorre notare le differenti scale di k per pioggia e neve: la seconda categoria è più soggetta all'azione del vento, per cui k può addirittura raddoppiare la quantità misurata per ottenere la quantità corretta, ovvero si stimano perdite fino al 50% sul totale.

La correzione si applica generalmente agli accumuli giornalieri, per precipitazioni di lieve-media intensità, poiché all'aumentare del tasso di precipitazione l'errore percentuale si riduce.

2.2.2. Modelli disponibili sul mercato

Le possibili tipologie di pluviometro accettate dalla WMO sono 3: a peso, a sifone e a bascula. Di questi tre, il primo è adatto a tutti i tipi di precipitazione, mentre gli altri sono principalmente atti a misurare pioggia liquida.

Il funzionamento è facilmente intuibile dal nome. Il pluviometro a peso misura la quantità caduta all'interno del collettore per differenza di peso rispetto alla quantità precedentemente contenuta.

Il pluviometro a sifone o a galleggiante sfrutta un contenitore in grado di riempirsi, in cui un galleggiante rileva la differenza di altezza della precipitazione raccolta.

Infine, il pluviometro a bascula sfrutta due contenitori che si riempiono e svuotano alternativamente ruotando su un perno al raggiungimento di un volume di acqua noto e proporzionale alla dimensione della bocca. Quest'ultimo è oggi il metodo più diffuso per le misure in remoto poiché non prevede un accumulo di acqua da rimuovere con regolarità in quanto si svuota automaticamente ogni quinto di mm raccolto. Siccome i pluviometri utilizzati in questo studio appartengono a questa categoria, si enunceranno di seguito le sue caratteristiche con maggior dettaglio.

Il pluviometro a bascula, di cui un esempio si trova in Figura 6 e in Figura 7, è l'ottimale per misurare l'intensità di precipitazione: ogni svuotamento indica una misura ben precisa ed è semplice risalire al volume di acqua precipitato. La bascula si ribalta in base al lato nel quale è localizzato il peso maggiore, ed il contenuto in massa d'acqua di ciascun contenitore è noto (m [g]). A seconda del modello questo volume è differente, ma è sempre proporzionale all'imbuto sovrastante il quale

convoglia l'acqua all'interno della bilancia, in modo che il volume corrisponda ad un'altezza di precipitazione di 0,2 mm. Supponendo costante la densità dell'acqua $\rho = 1 \text{ [g/cm}^3\text{]}$, si può stabilire una relazione massa-volume, e ricavare l'altezza h [mm] di precipitazione utilizzando l'area dell'imbecco S [cm²]:

$$V = \frac{m}{\rho} = h * S$$

1 mm corrisponde a 1 g su 10 cm², oppure più semplicemente, si intende 1 mm/m², che corrisponde al più comune 1 l/m², rendendo più semplice comprendere i volumi in gioco nonostante il litro non sia la misura volumetrica accettata nel Sistema Internazionale di unità di misura.



Figura 6 Pluviometro a bascula: è visibile il pendolo esterno alla bascula (con il magnete) e i contatti che permettono di inviare il segnale quando il reed switch viene attivato

Una volta rilevato lo scatto della bilancia però, occorre tenerne traccia. I primi modelli, adottati fino all'inizio degli anni 2000, sfruttavano un cilindro rotante ricoperto di carta. Un pennino era collegato al sistema basculante così che ad ogni svuotamento tracciasse un segno sulla carta. Il rullo aveva ovviamente una lunghezza finita e lo strumento richiedeva una manutenzione molto frequente, senza contare che i dati relativi alle precipitazioni potevano essere interpretati solo molto tempo dopo che la precipitazione era stata osservata. Dalla fine del secolo scorso e con sempre maggior sviluppo nei primi anni 2000 si è imposto un sistema elettronico basato sulla chiusura di un contatto: la doppia bilancia, visibile in Figura 7, è collegata ad un magnete che, muovendosi, ha effetto su delle lamelle

metalliche le quali chiudono un circuito permettendo di registrare una differenza di tensione (questo particolare tipo di sensore si chiama reed switch).

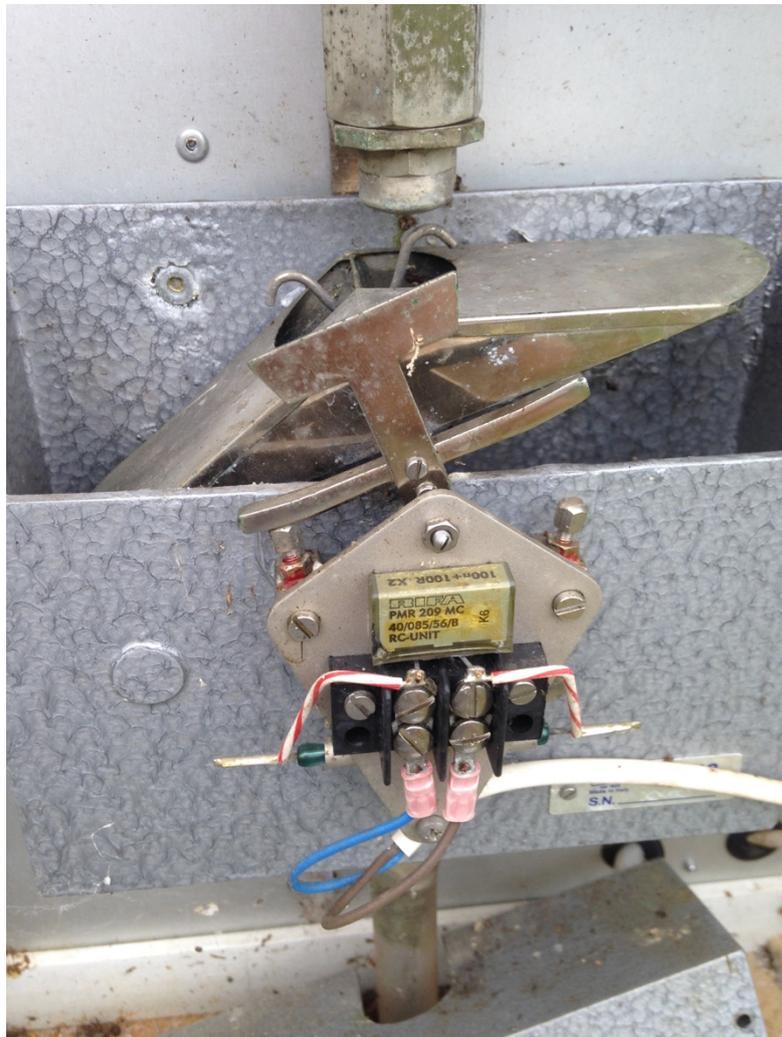


Figura 7 Doppia bascula inserita all'interno di un pluviometro, si distingue il meccanismo per convogliare l'acqua all'interno di una delle due vaschette in modo alternato, le quali sono chiuse per evitare un'eccessiva evaporazione. Alla bascula è collegato un magnete, al fondo di un piccolo contrappeso a forma di croce, il quale muovendosi insieme alla bascula attiva il sensore reed switch permettendo la registrazione di un quantitativo fisso di precipitazione proporzionale alla grandezza della bocca di ingresso e della vaschetta.

Nonostante sia stato inventato negli anni '40, non è stato utilizzato in questo ambito se non di recente. Le sottili lamelle si trovano all'interno di un piccolo cilindro di vetro, in atmosfera controllata e solitamente riempita con un gas inerte (solitamente azoto), e sono separate da pochi decimi di millimetro ma in posizione parzialmente sovrapposta, non permettendo usualmente il passaggio di corrente. Il magnete genera con il suo movimento a pendolo un campo magnetico che sposta le lamelle una contro l'altra permettendo il passaggio di corrente: l'impulso elettrico può essere registrato da un data logger o da un convertitore analogico-digitale e, in seguito, essere convertita in misura di pioggia, calcolata semplicemente contando il numero di ribaltamenti della bascula in 1 minuto. Questo sistema permette di registrare per lunghi periodi

senza che sia necessario l'intervento umano, nel caso in cui le misure siano inviate ad un database ad ampia capienza, evitando contemporaneamente di finire lo spazio di archiviazione su una memoria digitale locale.

Alcuni errori possono influire sulla correttezza del pluviometro a bascula: verranno elencati di seguito, congiuntamente alla relativa soluzione proposta da WMO ed osservata sul campo durante lo svolgimento della tesi:

- La bascula impiega un tempo finito per muoversi, seppur piccolo, durante la prima metà del quale altra acqua potrebbe entrare all'interno del contenitore già in fase di svuotamento. Questo errore meccanico, apprezzabile solamente durante piogge molto intense, è di tipo sistematico: si può ovviare con una precisa calibrazione dello strumento e utilizzando un sistema che faccia gocciolare a velocità controllata l'acqua. In Figura 6 si può notare un cilindro esagonale, circolare all'interno, posto al di sopra della cassetta metallica. Al fondo di esso vi è una strozzatura che aiuta l'acqua a gocciolare con regolarità, unitamente ad una guida metallica che ne permette l'orientamento verso la bascula attualmente in fase di riempiendo.
- La superficie dell'acqua raccolta all'interno della bascula è soggetta ad evaporazione. È sufficiente isolare in questo caso la bascula. Nel caso osservato essa consta di un doppio contenitore metallico, chiuso, che permette l'entrata e l'uscita dell'acqua da due piccole fessure. A sua volta l'intera bascula è all'interno di un contenitore metallico, isolato dall'esterno in una cabina di materiale plastico, bianco e altamente riflettente per evitare di surriscaldarsi.
- Durante piogge lievi la misura è soggetta ad uno spostamento nel tempo: occorre un periodo più lungo perché si riempia la vasca della bascula (Habib, Krajewski e Kruger 2001). Una possibile soluzione è di utilizzare uno strumento con risoluzione inferiore (0,1 mm), ma questo comporterebbe altri problemi di errore della misura. Oppure è possibile correggere la misura tramite una serie di parametri ausiliari e calcoli per spostare temporalmente la precipitazione (Colli, Lanza e La Barbera 2013).
- L'acqua può aderire alle pareti della bascula e modificarne la distribuzione dei pesi e il moto sul fulcro. Occorre pertanto scegliere un materiale adatto che limiti l'adesione. Un'altra alternativa consiste nel porre della cera sulle pareti: i test effettuati dal WMO hanno riportato una differenza del 4% di volume di acqua necessario per ottenere il ribaltamento, ovvero senza cera il 4% dell'acqua non si distacca dalle pareti. Questo problema non riguarda soltanto la bascula ma l'intero sistema collettore.
- La bascula è soggetta ad attrito ed alla forza di gravità, motivo per cui è molto importante che il pluviometro sia posizionato in posizione

perfettamente orizzontale. Spesso sul piano di appoggio del pluviometro o nei suoi pressi è posta una bolla circolare.

- Il sistema a bascula è altresì sensibile a movimenti differenti a seconda dell'intensità di pioggia. Questa problematica si risolve nuovamente tramite calibrazione e con test pratici di funzionalità.
- Altre correzioni pratiche sono suggerite nel documento di linee guida prodotto dalla WMO nel 2009, tra cui la principale è effettuare regolarmente una calibrazione dinamica inserendo un volume noto all'interno dello strumento, oppure applicando correzioni a livello di software sulla base delle curve suggerite (World Meteorological Organisation 2017), di cui un esempio sono le curve in Figura 5.
- In alternativa, WMO propone di condurre una procedura di modifica a livello di scheda elettronica del pluviometro per generare impulsi extra all'aumentare dell'intensità di precipitazione per non incorrere nella mancata misura. All'interno delle linee guida WMO, i risultati di questa procedura evidenziano un incremento delle prestazioni nella misura della precipitazione totale, ma errori nel tasso di intensità. Tuttavia, il metodo della correzione a posteriori tramite software ha mostrato risultati maggiormente positivi.

2.2.3. Rete locale utilizzata

Lo svolgimento di questo studio è stato reso possibile grazie alla collaborazione della 3a S.r.l., un'azienda sita a Torino ed attiva nel campo della meteorologia applicata all'agricoltura, fornendo ulteriori servizi di supporto e modelli previsionali fitopatologici.

Questa azienda ha infatti in gestione una rete pluviometrica, in parte di sua proprietà, di cui verranno enunciate a seguire le caratteristiche principali.

I primi albori dell'attività della 3a risalgono all'installazione delle prime centraline meteorologiche nel 1997, ma nel tempo trascorso fino ad oggi, il numero di stazioni è aumentato gradualmente ed oggi la piattaforma cloud della 3a gestisce i dati agrometeorologici di oltre 130 stazioni sul territorio piemontese, la cui distribuzione è visibile in Figura 17 al capitolo 3.3. Queste coprono l'intera area agricola della regione, ovvero la zona cerealicola, le colline adibite principalmente a colture viticole e corilicole (noccioleti) nel territorio patrimonio mondiale dell'Unesco di Langhe-Roero e Monferrato, la zona del cuneese a vocazione principalmente ortofrutticola e l'area di Vercelli, nota per le sue risaie. Nonostante la ridotta estensione, il Piemonte presenta una notevole varietà di colture e tipologie di terreno, motivo per il quale il monitoraggio meteorologico risulta essere ancora più importante.

Infatti, la Regione stessa fornisce agli agricoltori un servizio di monitoraggio a terra: tra le centraline in gestione alla 3a S.r.l. ve ne sono alcune di proprietà regionale. Questa frazione di stazioni fa parte della RAM (Regione Piemonte 2019), la Rete AgroMeteorologica del Piemonte. L'importanza di questa rete è ben espressa dalla Regione attraverso il suo stesso sito:

“Nel corso degli anni il servizio si è consolidato fino a diventare un supporto essenziale per il mondo agricolo e per la gestione dei processi produttivi. Anche le normative europee e nazionali sottolineano l'importanza fondamentale di questo tipo di supporti e prevedono che le Regioni assicurino la realizzazione di reti di monitoraggio e di sistemi di distribuzione delle informazioni.

La Rete piemontese consta attualmente di circa 120 stazioni meteorologiche di tipo elettronico di cui 63 sono di proprietà regionale mentre le altre sono di proprietà di enti pubblici o privati. Il funzionamento continuo nell'arco degli anni ha consentito l'alimentazione della banca dati che ormai ha assunto un valore storico di riferimento assai rilevante. Molte stazioni funzionano ininterrottamente ormai dalla seconda metà degli anni '90. In ogni caso la maggior parte delle stazioni possiede oramai una serie storica climatica di oltre 10-15 anni oggi preziosissima per analisi climatologiche capillari e interpretazione dei fenomeni atmosferici, agricoli e fitosanitari del presente rispetto al passato.”

Le stazioni di proprietà regionale sono tutte marchiate Siap-Micros, un'azienda italiana con sede a Treviso.

Delle stazioni citate, quelle di proprietà privata appartengono ad aziende agricole sperimentali, enti di ricerca oppure a privati. Tuttavia, la gestione dell'intero insieme delle stazioni è affidata alla 3a, con numerosi vantaggi in fatto di manutenzione e gestione unica del controllo qualità: tutti i parametri misurati dalle centraline sono inviati e registrati su un unico database organizzato su server cloud.

Prima di comprendere il sistema di gestione in remoto è opportuno spiegare la struttura della singola stazione e il suo funzionamento. Le centraline sono di tipologie differenti: quelle più moderne trasmettono i dati con cadenza oraria, le altre registrano i dati con la medesima frequenza, ma l'importazione nel database avviene a intervalli differenti, due o tre volte ogni 24 ore a seconda del modello.



Figura 8 Differenti modelli di stazioni meteo. Da notare in particolare la differente forma dei pluviometri (a sinistra un collettore con bocca da 1000cm^2 , a destra invece 500cm^2)

L'energia per alimentare la scheda elettronica e permettere l'invio è fornita da una batteria da 12V, la quale viene ricaricata tramite pannello solare. Le due procedure (registrazione e invio) sono indipendenti: ovvero è possibile che la stazione registri ma non invii i dati per problemi di rete. Tuttavia, è presente un data logger che conserva i dati delle registrazioni fino a 2-3 settimane, per poterli inviare quando possibile oppure, nel caso vi sia un guasto, finché l'operatore sopraggiunge a ripararlo ed ha la possibilità di scaricare i dati su un computer da campo, poiché inviare i dati di diverse settimane richiederebbe un tempo eccessivo ed un notevole dispendio di energia da parte della stazione. I dati sono infatti inviati grazie alla tecnologia GSM o GPRS (ovvero due differenti tecnologie di telefonia mobile cellulare a velocità limitata) sfruttando un modem. Bisogna considerare che le stazioni, seppur non tutte equipaggiate nello stesso modo, comprendo i seguenti sensori:

Tabella 1 Sensori e risoluzione degli strumenti installati su una stazione meteorologica della rete 3A

Sensore	Risoluzione	Unità
Pluviometro	0,2	mm
Termometro	0,1	°C
Umidità relativa	0,1	%
Umidità relativa del suolo	0,1	%
Anemometro	0,1	m/s
Direzione del vento	1	gradi
Bagnatura fogliare	1	min/ora
Temperatura del suolo	0,1	°C
Radiazione solare	1	MJ/m ²
Irradianza	1	W/m ²
Portata (alcune stazioni)	0,1	l/s

Per lo scopo di questa tesi, tuttavia, è importante analizzare nello specifico le caratteristiche dei pluviometri installati:

Tabella 2 Specifiche di un pluviometro installato su una stazione meteorologica della rete 3A

Risoluzione	0.2	mm
Accuratezza	3.0	%
Range operativo	0 ÷ 999.8	mm
Bocca	500/1000	cm ²

Il corpo del pluviometro è costituito da un cono aperto e verniciato con colori chiari per massimizzare la riflessione della radiazione solare incidente, con una superficie di 500 o 1000 cm² a seconda dei modelli (conforme a Guida 8 WMO, Cap 6 (World Meteorological Organisation 2017)). Il resto della struttura è costituito in materiale metallico altamente resistente agli agenti atmosferici, ai raggi UV, all'impatto e agli urti (es alluminio anodizzato). In Figura 8 si possono apprezzare due diversi modelli, i più comuni adottati nella rete pluviometrica considerata. La differente forma e area di cattura dei pluviometri coincide con una differente struttura interna del sistema basculante (ad esempio quella in Figura 6 appartiene al modello a sinistra, meno recente). Fanno parte delle stazioni pluviometriche in figura anche i sensori di umidità e temperatura (protetti dalla luce solare diretta da una serie di "cappelli" che danno un aspetto a fisarmonica al sensore); un rilevatore di radiazione solare incidente; un sensore elettrico per la simulazione della bagnatura fogliare; un anemometro e la banderuola per la direzione del vento.

2.2.4. Database e controllo di qualità

Le centraline meteo sono purtroppo soggette alle intemperie e all'usura, per cui richiedono una manutenzione periodica: ad esempio è comune osservare una caduta di tensione della batteria, per cui in seguito a prolungati periodi piovosi essa non è più in grado di ricaricarsi tramite il pannello solare; tutti gli altri sensori invece hanno una vita nominale compresa tra 4 e 6 anni, per cui periodicamente occorre cambiarli per ottenere una misura sempre precisa. Nonostante queste operazioni siano effettuate diligentemente, in alcuni casi i dati che giungono presso il database possono essere corrotti o mancanti in alcune parti: questo può accadere per molteplici motivi, dalla mancanza di segnale telefonico alla rottura di un sensore. In questi casi si rende necessario attuare una serie di procedure per garantire la qualità dei dati acquisiti in campo, tra cui ricostruire il dato sulla base di interpolazioni spaziali dei dati provenienti dalle stazioni vicine, comprendendo anche altri parametri ausiliari.

Al fine di garantire l'integrità della serie storica, infatti, il fornitore è tenuto a garantire e documentare le procedure per la ricostruzione automatica dei dati mancanti mediante interpolazione dei dati orari e giornalieri; le procedure devono ricostruire il dato con precisione sulla base dei dati delle stazioni comparative e della stazione stessa, utilizzando sia metodi di interpolazione spaziale (IDW) che di regressione lineare. Non è possibile, per motivi di segreto professionale, divulgare le formule utilizzate per l'interpolazione e i parametri adottati.

Si devono inoltre implementare procedure di ricostruzione per le seguenti grandezze meteorologiche: temperatura dell'aria, umidità dell'aria, precipitazione, bagnatura fogliare, radiazione globale giornaliera e velocità del vento giornaliera.

Per evitare una completa distorsione dei dati, oppure l'irreversibile perdita degli stessi, le procedure di correzione non devono sovrapporsi all'archivio dei dati originali, bensì generare un archivio operativo in cui i dati corretti o manipolati devono essere contrassegnati in maniera identificabile: si creano così 2 database, quello dei dati grezzi, e quello dei dati corretti ove necessario.

La valutazione della qualità dei dati deve essere realizzata secondo criteri riguardanti i singoli valori (coerenza della misura) e l'intera rete (confronto con dati di località vicine).

Tutte le procedure automatiche generano report quotidiani al termine e garantiscono la pubblicazione del dato, post controllo di qualità, entro le ore 8 del mattino.

L'attività di controllo ha l'obiettivo di evidenziare anomalie di funzionamento della base dati presente nell'archivio grezzo, derivanti anche dall'analisi automatica, come ad esempio valori tensioni di batteria, derive sulla misura dei sensori, irregolarità sulla trasmissione dei dati, intasamento di pluviometri, ecc.

Tali procedure manuali, insieme a quelle automatiche, consentiranno di importare il dato grezzo mancante (in caso di anomalia di trasmissione) o di ricostruire il dato anomalo e di individuare gli apparati in avaria per attivare l'intervento di manutenzione correttiva, anche in mancanza di una richiesta esplicita da parte dell'utente.

2.3. Satelliti

Strumento principe all'interno dell'ampia cornice del telerilevamento, ciò che comunemente definiamo satelliti sono più propriamente strumenti artificiali posti in orbita attorno al nostro pianeta. Essi permettono all'uomo di raggiungere un punto di vista differente, nel quale, grazie allo sviluppo delle comunicazioni e della trasmissione di dati è possibile utilizzare uno strumento in remoto.

Un satellite artificiale svolge una funzione principale in base ai dispositivi che trasporta, ovvero il "carico utile" o *payload*, ed una serie di funzioni ausiliarie necessarie per mantenerlo in orbita. Motori, combustibile, pannelli solari, antenne per la ricetrasmisione, protezioni, isolanti: sono i sistemi che debbono essere integrati nel satellite per mantenerlo in funzione e salvaguardare gli strumenti di misura primari.

A volte un singolo punto di osservazione non è sufficiente, oppure sono necessarie misure di diversa natura: si rendono quindi necessari gruppi di satelliti che possano agire in maniera corale. Questi gruppi sono detti "costellazioni".

Nell'ambito di questa tesi, il carico utile è rappresentato da sensori ottici. L'ottica è la scienza che studia il comportamento della luce e le sue interazioni con la materia: è, dunque, una branca dell'elettromagnetismo, data la natura duplice della luce di onda e particella.

L'unione delle scoperte nei campi della fotografia, dell'ottica umana e dell'elettronica ci permettono di indagare oltre il limitato campo del visibile. Lo studio dell'insieme multispettrale di radiazioni elettromagnetiche ci rende capaci di misurare e, quando possibile, sfruttare onde come i raggi cosmici, raggi gamma, X, ultravioletti, visibili, infrarossi, microonde e radio. Oltre a questo, la posizione dalla quale avviene l'osservazione di fenomeni non accessibili ai sensi umani possiede grossa rilevanza: fuori dall'atmosfera fenomeni di diversa natura e condizioni differenti possono favorire alcune misure e sfavorirne altre, basti pensare alla potenza dei moderni telescopi spaziali come Hubble.

In ambito elettromagnetico lo strumento che acquisisce i dati è detto *rivelatore*, mentre *sensore* è l'insieme di rivelatore e della struttura necessaria a rendere leggibile la misura, registrarla o trasmetterla.

La raccolta dall'atmosfera di dati riguardo a oggetti posti a distanza è detta *telerilevamento*, e può essere effettuata grazie alla misura di energia elettromagnetica. Quest'ultima può essere naturalmente emessa o riflessa dal corpo che si intende indagare, si parla in tal caso di misure passive, oppure è possibile misurare il risultato dell'interazione con l'oggetto di un'energia emessa da una fonte artificiale nota, regolabile a piacimento, quindi si effettua una misura attiva: si dice che il corpo è illuminato dalla fonte. Un radar, ad esempio, è un sistema attivo, mentre una misura passiva è solitamente quella ad infrarossi o

microonde. Volendo fare un paragone, è come considerare una macchina fotografica con o senza flash.

Le fasi critiche del telerilevamento sono 3:

- Acquisizione dei dati
- Elaborazione dei dati
- Interpretazione e utilizzo

All'interno di questo studio verrà spiegato in seguito e con maggior dettaglio quando queste fasi sono applicate.

Per ora ci limiteremo a elencare i principali vantaggi di un sistema di telerilevamento satellitare.

- L'osservatore è in posizione privilegiata: dall'alto si amplia lo spazio visibile da un singolo punto, non vi sono vincoli geopolitici né frontiere a limitare l'osservazione.
- Le informazioni sulla larga porzione di territorio inquadrata di volta in volta dallo strumento sono coerenti, in quanto provenienti dallo stesso strumento. Inoltre, in genere sono informazioni di carattere numerico, rendendo più usufruibile il dato e più semplice la manipolazione e il trattamento.
- La natura delle misure ottiche ne permette la registrazione in forma di pixel e griglie, permettendo automazioni e il conseguente accesso alle informazioni in tempo quasi reale (*near-real-time*).
- Risoluzione spaziale e temporale sono spesso correlati e permettono di effettuare analisi differenti privilegiando la prima o la seconda.

La risoluzione è un elemento importante dei sistemi di telerilevamento. Il concetto di risoluzione, infatti, riguarda il livello di dettaglio che è possibile raggiungere su differenti aspetti: spaziale, temporale e spettrale. Nel caso dei satelliti, la risoluzione spaziale è strettamente correlata allo strumento di acquisizione della radiazione, alla sua capacità di cogliere e distinguere oggetti relativamente vicini o lontani e di registrarne la radiazione emessa, sia essa la luce nel campo del visibile come in una fotografia o la *brightness temperature* nel campo delle microonde, ovvero la temperatura di corpo nero.

Se un'immagine appare a quadretti grossolani, i bordi degli oggetti sono frastagliati, le linee rette subiscono brusche interruzioni e addirittura alcuni elementi non vengono rappresentati, significa che la risoluzione spaziale è ridotta. Viceversa, un'alta risoluzione permette di rappresentare un oggetto in modo più simile alla realtà: in Figura 9 è disponibile un esempio di come la variazione della risoluzione comprometta in maniera sostanziale la nostra capacità di riconoscere un elemento o un colore (in concomitanza con la risoluzione spettrale).

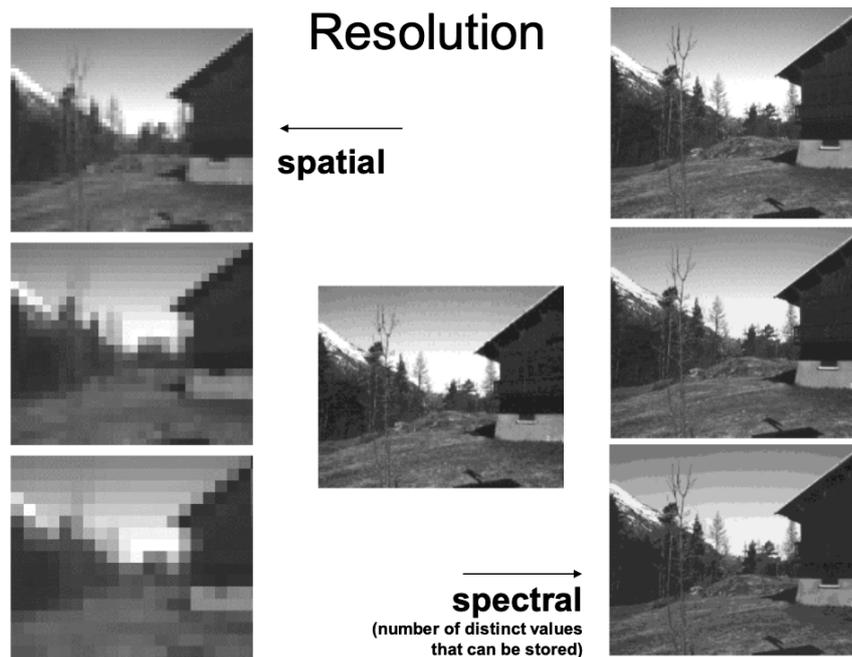


Figura 9 Esempio di variazione della risoluzione (Dispense corso GIS 2018, Técnico Lisboa). Una risoluzione spaziale elevata permette un miglior riconoscimento delle forme degli oggetti, ottenendo un maggior dettaglio. La risoluzione spettrale, rappresentata a destra, è legata alla capacità di distinguere valori differenti di una grandezza: in questa rappresentazione si tratta di bande di colore, nel caso oggetto di studio invece, si tratta di considerare intervalli di ampiezza più o meno elevata di lunghezza d'onda delle radiazioni elettromagnetiche.

La risoluzione temporale invece, è legata al numero di volte in uno stesso oggetto è inquadrato in un tempo definito. Si tratta di definire, in altre parole, con quale intervallo un satellite (o un satellite della stessa costellazione con il medesimo carico utile) transita al di sopra dell'oggetto. Un'elevata risoluzione temporale ci permette di osservare l'evoluzione di un fenomeno soggetto a cambiamenti nel tempo, ad esempio lo spostamento di un fronte nuvoloso, il dilagare di un incendio, l'avanzata della desertificazione, la crescita di una pianta, ecc.

Ovviamente una risoluzione migliore, sia temporale che spaziale, è da prediligere, ma ad un loro miglioramento corrisponde un proporzionale incremento dei costi. Per questo motivo, ad esempio, i satelliti a fini meteorologici hanno spesso una risoluzione spaziale minima ma acquisiscono numerose immagini con intervalli sub-orari.

Infine, un ulteriore aspetto degno di importanza, seppur di non immediata evidenza, è la risoluzione spettrale. Come anticipato, la quantità acquisita dagli strumenti satellitari in gioco in questo studio è la radiazione elettromagnetica.

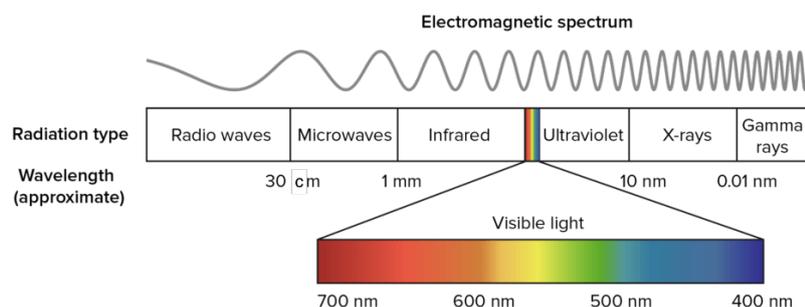


Figura 10 Spettro elettromagnetico. La stima di precipitazione considerata in questo studio è ricavata dalla misura della radiazione con lunghezza d'onda contenuta nelle bande delle microonde e degli infrarossi (Luca Fornaciari Fotografia s.d.)

Lo spettro elettromagnetico in Figura 10 rappresenta l'insieme di tutte le frequenze elettromagnetiche. Universalmente si adottano degli intervalli in cui dividere lo spettro, per cui in ognuno di essi le caratteristiche di energia sono simili.

Un rilevatore non è in grado di raccogliere tutte le lunghezze d'onda dello spettro, è limitato dalla sua risoluzione e in un determinato campo d'azione definito dal produttore. La risoluzione spettrale dunque rappresenta il livello di dettaglio raggiungibile nel distinguere le diverse lunghezze d'onda.

Per spiegare meglio il concetto consideriamo l'occhio umano. Questo è in grado di recepire raggi con lunghezza d'onda λ tra 390 e 760nm, e di distinguere milioni di colori, ognuno corrispondente ad una particolare banda di frequenza. Eppure, un intervallo qualsiasi è divisibile all'infinito, perciò ogni singola variazione di lunghezza d'onda dovrebbe corrispondere ad una percezione differente, ma l'occhio accomuna lunghezze d'onda vicine in bande, o colori. Non riusciamo a distinguere infiniti colori: allo stesso modo un rilevatore distingue un limitato numero di frequenze. Tornando ai concetti di risoluzione visti in precedenza, l'occhio possiede una risoluzione spaziale di qualche micrometro da vicino: due fili d'erba posti troppo vicino ci sembreranno uno solo. Oppure un oggetto in moto sembrare immobile o muoversi in maniera anomala: è quello che accade quando osserviamo il cerchione della ruota di un'auto in movimento o le pale di un elicottero, il cui movimento ripetuto può essere caratterizzato da una frequenza di rotazione troppo veloce per poter essere colta dall'occhio umano.

Questi 3 diversi aspetti della risoluzione rappresentano una notevole sfida tecnologica nella costruzione di un nuovo strumento di cattura su un satellite e dei risultati che si possono raggiungere con il suo lancio in orbita.

I principali strumenti satellitari agiscono nelle bande microonde (MW o *microwave*) o infrarossi (IR). Le microonde hanno lunghezze d'onda dell'ordine dei centimetri, per cui possono filtrare oltre le nuvole, a differenza gli IR, caratterizzati da lunghezze d'onda fino a mille volte più piccole. Tuttavia, la minor lunghezza d'onda permette agli infrarossi di avere una risoluzione spaziale maggiore delle

MW. Per questo motivo gli strumenti a microonde si trovano al di sopra di satelliti orbitanti a quote inferiori (350-850 km) e gli IR orbitano attorno ai 36000 km (Remote Sensing Systems s.d.).

La capacità di distinzione di elementi differenti è legata anche alla distanza a cui è posto l'oggetto di indagine e dalla capacità dello strumento di metterlo a fuoco. Per questo motivo disporre di più punti di osservazione per uno stesso fenomeno porta effetti positivi, come, in questo caso, l'utilizzo delle due fonti quali la rete di pluviometri della società 3a S.r.l. e il progetto GPM (*Global Precipitation Measurement*).

2.3.1. GPM

La missione Global Precipitation Measurement (NASA 2015) fa seguito ad un'altra del tutto simile: la TRMM (*Tropical Rainfall Measuring Mission*) (NASA 2015), una missione di studio dei fenomeni di precipitazione, fu lanciata nel 1997 come progetto congiunto tra NASA (*National Aeronautics and Space Administration*) e JAXA (*Japan Aerospace Exploration Agency*). Il progetto includeva un satellite il cui scopo era migliorare la capacità di cogliere la variabilità delle precipitazioni nella zona compresa tra i due tropici, e di poter meglio studiare il ciclo dell'acqua con una fonte di informazione quasi globale. La richiesta sempre in crescita di informazioni sui cicli che governano la distribuzione del calore, la circolazione di vapori, nuvole e precipitazioni di ogni natura fu il motore che spinse gli ideatori del progetto, oltre al desiderio di porsi in una posizione avanguardista nei confronti del cambiamento climatico. Infatti, per ottenere una qualità riconosciuta nello studio di fenomeni meteorologici caratterizzati una scala tanto ampia è fondamentale adottare un punto di vista universale. L'energia latente che il vapore atmosferico rilascia nella trasformazione in precipitazione liquida o che il passaggio inverso richiede per la formazione degli ammassi nuvolosi rappresenta infatti la più grossa entità in moto nei cieli. In generale, non è possibile misurare direttamente questa energia, ma è possibile ricavarla considerando la forte correlazione con la quantità di acqua in circolo in atmosfera, la maggior parte della quale si trova nella zona più calda del pianeta.

Dopo circa 17 anni dal lancio della missione TRMM, l'evoluzione della tecnica e la maggior attenzione rivolta agli studi climatici hanno portato i due enti spaziali, americano e giapponese, a sviluppare un progetto che si pone in perfetta continuità con il TRMM: il Global Precipitation Measurement o GPM.

Il nucleo della missione è rappresentato dal *GPM Core Observatory*, un satellite sviluppato appositamente da NASA e JAXA, che opera tramite i due principali sensori GMI (sensore microonde) e DPR (radar) le cui caratteristiche saranno analizzate con maggior dettaglio in seguito.

Il sistema GPM si compone tuttavia di numerosi elementi. Prima di tutto la struttura fisica di raccolta dati è costituita da una costellazione di satelliti di diversa origine, equipaggiati in maniera differente. Alla missione collaborano un nutrito gruppo di agenzie spaziali grazie ad accordi bilaterali: CNES (*Centre National d'Études Spatiales*), ISRO (*Indian Space Research Organization*), NOAA (*National Oceanic and Atmospheric Administration*), DOD (*Department Of Defence* degli USA), EUMETSAT (*European Organisation for Exploration of Meteorological Satellites*) ed altri ancora.

La costellazione internazionale copre l'intera superficie terrestre ogni 2-3 ore: infatti, sfruttando satelliti già in orbita, la frequenza di campionamento aumenta notevolmente (Figura 11).

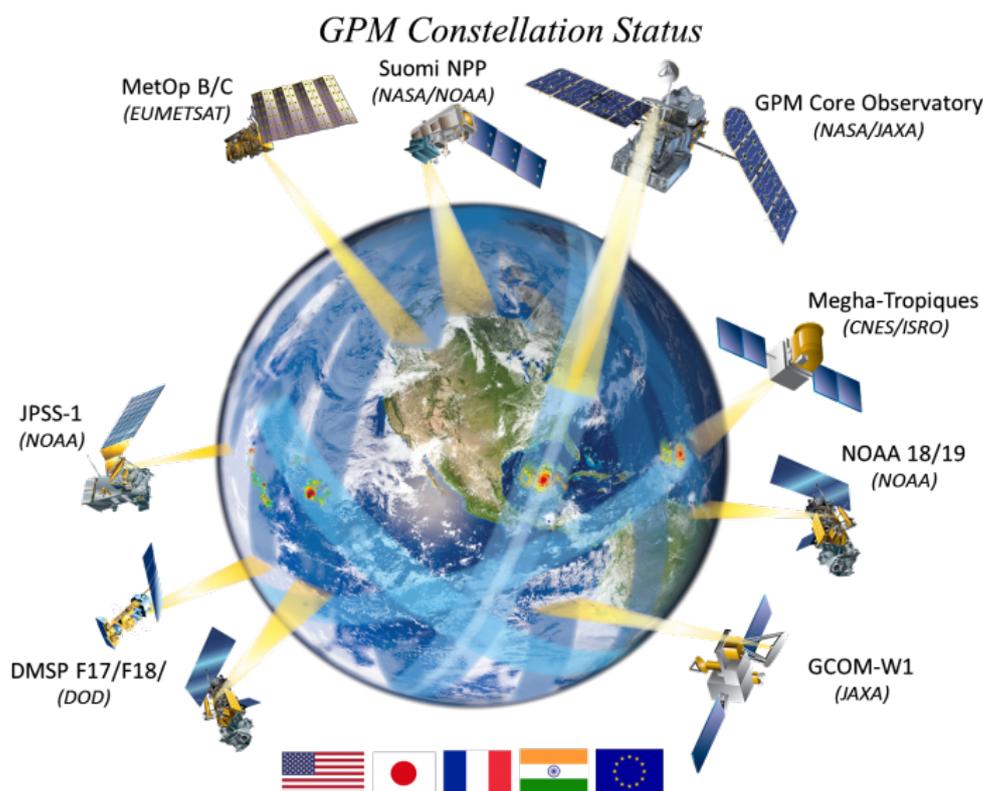


Figura 11 Schema della costellazione GPM, Satelliti principali e bandiere delle nazioni partner (NASA 2015) .

Il GPM Core Observatory, nonostante sia stato sviluppato e testato al Goddard Space Flight Centre, è stato lanciato il 27 Febbraio del 2014 da Tanegashima, in Giappone, sfruttando la spinta di un razzo di matrice nipponica H-IIA. Esso segue un'orbita sospesa a 407 km dalla superficie terrestre, in moto non eliosincrono, coprendo l'area tra 65°S e 65°N, ampliando l'estensione del suo predecessore TRMM. Inoltre, l'inclinazione del Core Satellite rispetto alla superficie è stata aumentata da 35 a 65°. L'esperienza della missione precedente ha permesso di comprendere quali elementi fosse necessario migliorare, producendo un nuovo satellite è molto più sensibile a neve e piogge deboli (anche < 0.5 mm/h), siccome a latitudini maggiori la loro frequenza aumenta rispetto alle intense precipitazioni

equatoriali. Per adempiere a questo nuovo obiettivo la strumentazione a bordo non è stata semplicemente migliorata, bensì variata.

2.3.2. Strumentazione GPM

A bordo del satellite cardine della missione si trovano due strumenti principali (NASA 2015):

- Dual-frequency Precipitation Radar (DPR): è composto da due radar attivi su due bande radiometriche differenti sono allineati in modo da fornire la stessa griglia di osservazione di 5km al suolo. Il primo, KuPR (Ku-band Precipitation Radar) è un'evoluzione del radar già presente sul satellite TRMM, ed opera a 13.6 GHz, mentre il secondo KaPR (Ka-band Precipitation Radar) agisce nella banda superiore delle frequenze a microonde (35.5 GHz). Il DPR, progettato appositamente per un utilizzo aerospaziale, primo in questo campo, è stato migliorato rispetto al suo equivalente della missione precedente per poter cogliere con più precisione pioggia di lieve entità, giungendo ad una sensibilità inferiore a 0,2 mm/h e potendo registrare su 3 dimensioni. Insieme al GMI forma il gruppo degli strumenti principali del corredo del GPM Core: sono in grado di lavorare di giorno e di notte, sia su terreno solido che sopra mari ed oceani. L'ampiezza di registrazione è tale per cui il KuPR scansiona ad ogni acquisizione una striscia di 245km e il KaPR 120km.

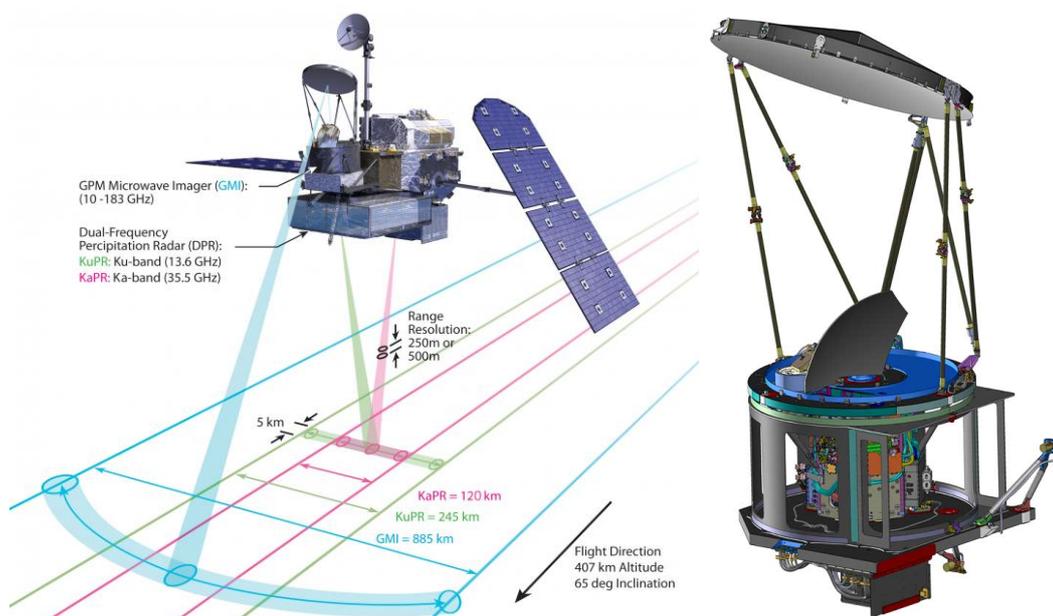


Figura 12 Rappresentazione del satellite GPM Core Observatory, con indicazione degli strumenti e del campo di azione: il sistema radar agisce tramite "spazzate" con ampiezza 120 o 245 km a terra, mentre il radiometro a microonde con una ampiezza di 885 km. È inoltre indicata la direzione e la velocità di moto e la posizione degli strumenti principali. A destra un ingrandimento del GMI

- Global Precipitation Measurement Microwave Imager (GMI): è un radiometro a microonde multi-canale e a scansione di forma conica. È fondamentale per poter avere copertura spaziale e temporale ad alta risoluzione. Infatti, il GMI è in grado di registrare la radiazione su 13 canali microonde (da 10 GHz a 183 GHz) spaziando quindi tra diverse lunghezze d'onda sub-metriche. Il set di frequenze è stato ottimizzato in seguito all'esperienza di telerilevamento degli anni passati in modo da registrare efficacemente precipitazioni di ogni intensità e sfruttando tutte le bande e le loro differenze radiometriche.

Possiede, rispetto al predecessore TMI montato sul TRMM, un'antenna da 1,2m di diametro che ha permesso un notevole miglioramento della risoluzione spaziale.

Il GMI "fotografa" un cono sulla superficie terrestre con un'incidenza tra normale alla superficie terrestre e superficie esterna del cono di $52,8^\circ$ e ruota su se stesso 32 volte al minuto: in questo modo misura la radiazione su una porzione di 885 km di diametro, centrata sulla normale al satellite. Inoltre, quest'area è centrata rispetto alle misure del DPR, con una conseguente sovrapposizione dati raccolti e un ulteriore miglioramento delle misure. Il GMI rappresenta quindi lo standard di precipitazione per i radiometri installati sugli altri satelliti della costellazione.

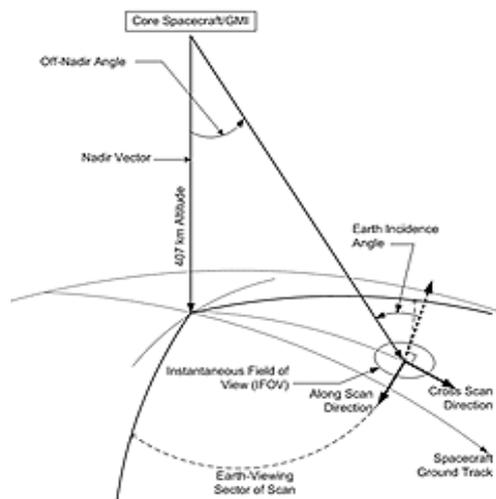


Figura 13 Diagramma dello scanning conico del GMI: Ruotando su se stesso ad un'altezza di 407 km dal suolo, il radiometro in quadra ad ogni rotazione l'area circoscritta dalla linea tratteggiata, composta da numerose misure istantanee circolari incidenti a terra con un angolo di $52,8^\circ$. Il radiometro effettua un giro completo 32 volte ogni minuto (NASA 2015).

Per evidenziare chiaramente la necessità di disporre di questi due strumenti si riassume il concetto principale: il DPR sonda i dati di precipitazione in atmosfera su 3 dimensioni, mentre il GMI cattura l'intensità di precipitazione sulla superficie terrestre e la sua distribuzione orizzontale. L'azione combinata di questi due strumenti è un valido riferimento per gli altri satelliti della costellazione i cui

parametri raccolti possono essere uniti per formare un dataset di precipitazione a scala globale.

Per poter funzionare, il GPM necessita di altre strutture atte al sostentamento energetico, alla trasmissione di dati, al mantenimento dell'orbita:

- SADDs: Solar Array Drive and Deployment Subsystem. Il *solar array* è l'unione di diversi pannelli solari, e questo è mosso e direzionato dal SADDs.
- HGAS: High Gain Antenna Subsystem. È il sistema di orientamento dell'antenna di trasmissione.
- TCS: Thermal Control Subsystem. Controlla la temperatura in modo che DPR e GMI funzionino in modo ottimale
- Propulsion Subsystem. È il sistema che si occupa di mantenere in orbita il satellite, correggendo le anomalie durante il funzionamento ma soprattutto nel momento di distacco dal modulo di lancio e durante il rientro in atmosfera alla fine del periodo di servizio.
- Power Subsystem. Controlla l'ingresso di energia dai pannelli solari, l'unica fonte di ricarica della batteria, e distribuisce l'energia a tutti gli altri strumenti.
- GN&C: Guidance, Navigation and Control Subsystem. Si occupa dei calcoli per mantenere l'inclinazione ottimale nei confronti della superficie terrestre nei limiti di accuratezza consentiti.
- RFComm: Radio Frequency Communication Subsystem. Fondamentale per la comunicazione terra-spazio, gestisce il sistema radio in banda S (2-4 GHz). Oltre alla trasmissione dei dati misurati, le trasmissioni includono resoconti periodici sullo stato generico del satellite, trasmissioni di emergenza e reportistica di ogni sistema e sottosistema.
- C&DH: Command and Data Handling Subsystem. Immediatamente prima della trasmissione dei dati, il sistema C&DH raccoglie e processa le informazioni sulle tutte le componenti che possono essere monitorate in remoto
- SMAC: Safety Mechanisms and Attitude Control Subsystem. Provvede a generare segnali di controllo e informazioni di stato sui differenti sensori e sistemi, inclusi quelli sopracitati, e dai propulsori. Sotto a questo sistema vi sono ulteriori meccanismi indipendenti di controllo per evitare che un guasto possa danneggiare l'intero satellite.
- Flight Software. La parte informatica del satellite include una moltitudine di programmi, procedure, algoritmi e procedure che controllano il funzionamento corale del satellite e dirigono le sue operazioni. (NASA 2015)

Nonostante ciò, il GPM Core Observatory non sarebbe in grado di raggiungere un'elevata risoluzione temporale da solo. Per questo motivo vi sono altri sensori basati sulle frequenze del microonde, sia attivi che passivi, che contribuiscono a raggiungere l'obiettivo. Verranno elencati di seguito i principali:

- Special Sensor Microwave Imager/Sounder (SSMIS) sul satellite DMSP dello U.S. Defense Meteorological Satellite Program.
- Advanced Microwave Scanning Radiometer-2 (AMSR-2) della missione JAXA Global Change Observation Mission - Water 1 (GCOM-W1)
- Multi-Frequency Microwave Scanning Radiometer (MADRAS) e il multi-channel microwave humidity sounder (SAPHIR) sul satellite Megha-Tropiques fornito dalla francese CNES (Centre National D'Etudes Spatiales) e dall'indiano Space Research Organisation (ISRO)
- Microwave Humidity Sounder (MHS) all'interno della rete NOAA
- MHS montato sui satelliti della serie MetOp, progettati dal progetto europeo EUMETSAT
- Advanced Technology Microwave Sounder (ATMS), incluso nel National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP), anch'egli frutto di un progetto americano
- ATMS, un ulteriore figlio della collaborazione NOAA-NASA per la missione Joint Polar Satellite System (JPSS).

2.3.3. Segmento a terra ed algoritmi di calcolo

La misura e la trasmissione dei dati non sono le uniche operazioni necessarie alla produzione un dato usufruibile: è fondamentale una struttura elettronica e informatica alla base che sia in grado di processare una notevole quantità di byte di dati. Si tratta quindi di costruire un algoritmo efficace: all'interno della missione GPM esso è denominato IMERG: *Integrated Multi satellitE Retrievals for Gpm*. La spiegazione più chiara delle capacità di questo algoritmo si trova nelle note tecniche della missione GPM stessa:

"This algorithm is intended to intercalibrate, merge, and interpolate "all" satellite microwave precipitation estimates, together with microwave-calibrated infrared (IR) satellite estimates, precipitation gauge analyses, and potentially other precipitation estimators at fine time and space scales for the TRMM and GPM eras over the entire globe." (Nasa Algorithm Theoretical Basis Document V06 2019)

Segue la traduzione:

“L’algoritmo intende intercalibrare, unire e interpolare “tutte” le stime di precipitazione effettuate tramite microonde, insieme alle stime basate sugli infrarossi, calibrate tramite microonde, le analisi provenienti dalle reti pluviometriche a terra e, potenzialmente, altri strumenti di stima ad elevata risoluzione spaziale e temporale per il TRMM e il GPM.”

Il fine ultimo di IMERG è ambizioso: unire tutte le sorgenti possibili per crearne una universale e capace di unire gli aspetti positivi di ogni elemento. L’algoritmo si nutre di numerose fonti, per restituire un’elaborazione di dati per ogni mezz’ora, distribuiti su di una griglia con ampiezza $0,1 \times 0,1^\circ$. Tuttavia, è necessario un tempo di 2-3 ore per poter ottenere un primo risultato: questo intervallo è inoltre da sommare al tempo necessario per ricevere i dati dai diversi satelliti a livello mondiale e successivamente rendere pubblico l’output. Il risultato è quindi consultabile con 4 ore di ritardo rispetto al momento in cui è effettuata ciascuna misura. IMERG è strutturato in modo da poter essere applicato numerose volte: sono quindi disponibili all’utente 3 diversi risultati in base al tempo di elaborazione, ovvero a quante volte l’algoritmo ha potuto operare: *Early*, *Late* e *Final Run* sono i nomi degli output, disponibili rispettivamente dopo 4 ore, 12 ore e 3 mesi e mezzo. La differenza elevata di tempo tra Late e Final run è dettata dal fatto che la seconda elaborazione integra anche dati provenienti delle reti pluviometriche, le quali però non sono gestite direttamente all’interno della missione.

Infine, i dati sono passati al PPS (*Precipitation Processing System*), ereditato dall’equivalente TSDIS (*TRMM Science Data Information System*) della missione precedente. Il PPS si occupa di processare e archiviare i dati provenienti da IMERG e dagli altri algoritmi secondari del GPM. Proprio tramite il PPS è possibile scaricare e consultare i risultati del progetto, includendo anche output grafici utilizzabili in software di elaborazione tipo GIS, ma anche immagini e grafici a fini divulgativi.

I meccanismi interni che permettono all’algoritmo di generare una stima di precipitazione a partire da misure radiometriche non sono divulgati dai loro creatori, ma è possibile intuire i passaggi fondamentali del processo. Le misure di tipo *passive microwave* sono effettuate soltanto da satelliti a “bassa” quota (*low Earth orbit* o *leo*) come il Core Observatory, perciò il cono di osservazione è limitato: basta guardare alle specifiche del satellite GPM per rendersi conto che un’inquadratura di circa 900 km non è sufficiente. Per questo motivo, l’algoritmo è studiato per unire diverse “fotografie” satellitari in un mosaico. Infatti, è in grado di riempire lo spazio di osservazione tramite interpolazione lineare, e dove ancora vi è una carenza di informazioni a causa di immagini troppo distanziate, IMERG colma questo spazio con stime derivate da osservazioni a infrarossi elaborate tramite un filtro Kalman, in particolar modo a latitudini superiori a 65° dove il GPM Core Observatory non è in grado di spingersi.

A fini di ricerca sono disponibili anche output intermedi (livello 1 e 2) rispetto al risultato finale di IMERG, che rappresenta il terzo e ultimo livello di elaborazione. Il livello 1 a sua volta contiene 3 sottolivelli, di cui il primo (detto 1A) contiene i dati di radiazione non ancora processati, con un minimo livello di trasformazione, unendo informazioni su efemeridi e coefficienti di calibrazione, giungendo infine al livello 1C, composto dalla *brightness temperature*, quindi ancora a livello di radiazione. Il livello 2 rappresenta una prima stima di precipitazione proveniente dai soli strumenti DPR e GMI del Core Observatory: è compito finale di IMERG unire queste osservazioni a quelle provenienti dal resto della costellazione, il cui risultato è distribuito sotto forma di livello 3. Questo tipo di elaborazione a tappe è ripetuta un numero di volte sempre maggiore per Early, Late e Final Run. Occorre prestare particolare attenzione al risultato detto Final Run, il quale si presenta come il livello più elevato di elaborazione e di inclusione di dati con origini differenti.

Una volta raccolti i dati da parte della costellazione alla base operativa GPM viene creato il mosaico di unione, poi calibrato nel livello 1, convertito in precipitazione nel livello 2 e, successivamente, diviso in settori in una griglia e diviso in stime 30-minutali di intensità di precipitazione tramite una moltitudine di processi e algoritmi secondari (GPROF, CORRA, MERRA, PERSIANN, MORPHING KALMAN FILTER, quasi-lagrangian time interpolation). Inoltre, nella versione Final, le precipitazioni stimate sono calibrate con coefficienti climatologici che variano in base al periodo dell'anno e alla posizione geografica. Perché questo accada nel migliore dei modi, il giorno da calibrare viene mantenuto al centro del periodo considerato per la calibrazione, in modo che i coefficienti applicati siano una media tra quelli dei giorni precedenti e successivi. L'ultimo passaggio dell'algoritmo consiste nell'unire i dati provenienti da una rete pluviometrica mondiale: sono utilizzate circa 116'000 stazioni, la maggior parte delle quali sono incluse in un altro progetto tedesco. Il GPCC (*Global Precipitation Climatology Centre*) del Deutscher Wetterdienst (Deutscher Wetterdienst s.d.), ovvero il servizio meteorologico tedesco, si occupa anche della correzione delle misure pluviometriche per eliminare sovra e sottostime. Queste misure sono utilizzate per correggere lo scarto dei dati orari e mensile, in quanto all'interno della Final Run è presente anche un'aggregazione mensile derivata dalla somma dei dati orari.

3. Reperibilità dati e procedure applicate

3.1. Scelta e reperibilità dei dati

All'interno di questo studio sono state utilizzate due principali fonti di dati orari: quelli forniti dalla 3a S.r.l. si trovano in forma tabellare, sotto forma di file csv (*comma separated values*), uno per ogni anno, mentre i valori provenienti dal programma GPM-IMERG sono scaricabili dal sito ufficiale della missione in formato ".h5" o ".hdf5" (*Hierarchical Data Format*).

Il periodo che si è scelto di adottare comprende la fascia di anni tra il 2013 e il 2017 inclusi. All'epoca dell'inizio di questo studio, il 2019 non si era ancora concluso, per cui è stato escluso a priori. Tuttavia, il principale motivo della scelta dipende in larga parte dalla fonte GPM: infatti nel febbraio 2014 fu lanciato in orbita il GPM Core Observatory a sostituire e migliorare il precedente satellite TRMM, ma i dati precedenti al 2014 sono stati riprocessati nel 2019 con l'algoritmo IMERG. Nonostante l'utilizzo di una diversa fonte, i dati hanno subito lo stesso tipo di processing. Risulta quindi interessante indagare anche la bontà dell'algoritmo nel gestire dati non nativi, ovvero in questo caso i dati raccolti nel biennio 2013-2014. In seguito, il periodo tra il 2015 e il 2017 rappresenta un lasso di tempo sufficiente per trarre conclusioni ai fini di questa tesi.

Come visibile in Figura 14, la disponibilità di entrambe le fonti è continua. Eppure, dal grafico non traspare come nel tempo la qualità dei dati sia migliorata. Dal lato della rete pluviometrica è aumentato il numero di stazioni, passando da meno di 120 nel 2013 a 137 nel 2017; mentre per quanto riguarda la missione GPM, occorre precisare che il periodo tra la seconda metà del 2014 e l'8 aprile 2015 (giorno di spegnimento degli strumenti TRMM) fu un periodo di transizione tra strumenti di misura differenti, ma accomunati dallo stesso algoritmo di analisi (e ri-analisi).

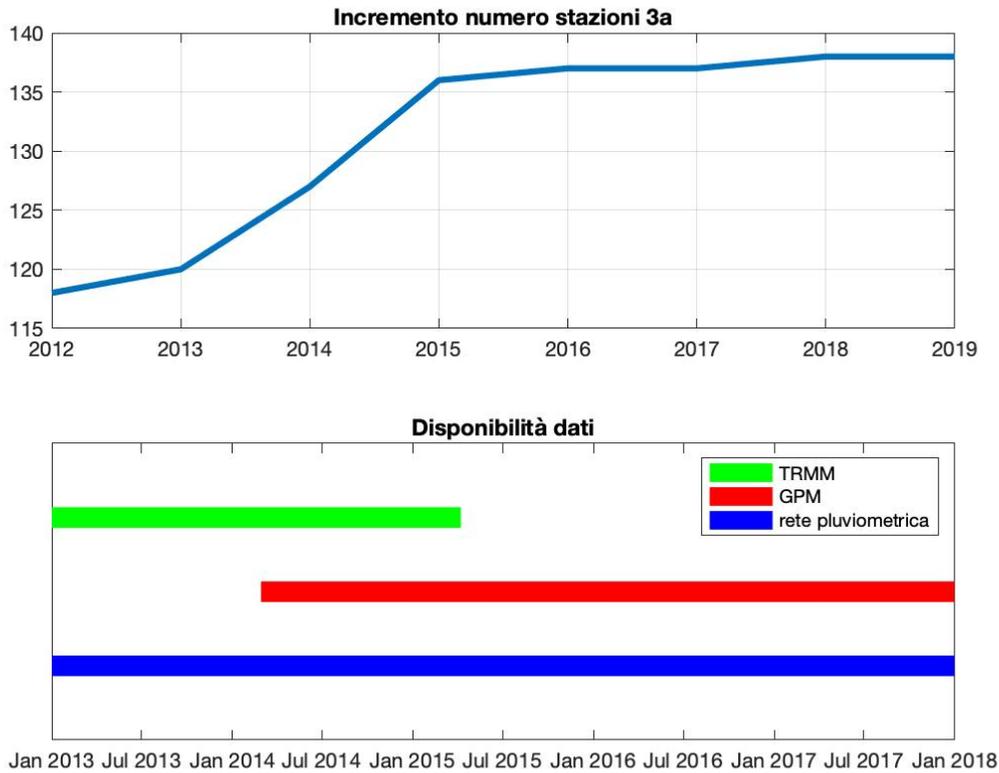


Figura 14 Sopra: rappresentazione del numero di stazioni in gestione alla 3a durante gli anni. Sotto: la disponibilità temporale dei dati della rete pluviometrica e quelli processati da IMERG. Fino a metà 2015 IMERG ha riprocessato dati provenienti dalla missione TRMM, e successivamente quelli raccolti dalla missione GPM

Vi sono numerose fonti di dati meteorologici, gratuite o meno, che è possibile usare ai fini di uno studio inerente la precipitazione: reti pluviometriche nazionali (Nashwan 2019) (Chen 2016) (Asong 2017); reti appartenenti a enti militari, quali l'aeronautica; è altresì possibile considerare le numerose singole serie storiche fornite dai numerosi aeroporti, dagli istituti di ricerca (Arpa, ad esempio), dalle università, protezione civile, ecc...il motivo per il quale la scelta è ricaduta su una rete pluviometrica a scopi agrometeorologici da un lato e la missione GPM dall'altro, risiede nell'opportunità di utilizzare una rete a terra le cui stazioni sono caratterizzate da una distribuzione molto capillare sul territorio. Infatti, numerosi studi che ambivano a studiare l'efficacia del GPM hanno utilizzato fonti a scala nazionale, in cui le reti pluviometriche erano spesso costituite da poche stazioni su un'area molto ampia. Inoltre, la maggior parte di questi studi ha operato un confronto su un'area tendenzialmente pianeggiante. Vi sono alcune eccezioni (Caracciolo 2018), tuttavia il territorio piemontese presenta condizioni uniche e scarsamente rintracciabili in altre aree del pianeta: sono infatti presenti rilievi montuosi molto elevati, il larga parte sopra i 3000m, disposti lungo un arco alpino di notevole estensione, e contemporaneamente la distanza dal Mar Mediterraneo è ridotta, motivo per il quale le pendenze dei declivi possono essere molto elevate.

In seguito ad un tirocinio curricolare presso la 3a S.r.l., è sorta l'opportunità di utilizzare una rete pluviometrica ad alta densità, per cui il confronto tra le due fonti di dati pluviometrici sopracitate è stato dettato dai seguenti fattori:

- Non sono presenti studi di validazione della missione GPM nel nord Italia;
- La conformazione orografica della Regione è molto particolare, se non unica, in Italia e in Europa: il 48% del territorio infatti è montuoso, il 20% collinare e solo il 32% pianura, come visibile in Figura 15. Inoltre, si tratta di un territorio di modesta estensione costituito da una zona di pianura circondata dai rilievi alpini e diffusi sollevamenti collinari nella zona centrale. Per questo motivo, anche se la totalità dei pluviometri è al di sotto di 820 metri di altitudine e, guardando in Figura 27, si nota che sono principalmente sotto i 500 metri, è interessante studiare l'applicabilità della stima della missione GPM.
- Studi precedenti hanno considerato una rete pluviometrica puntuale molto rada (Nashwan 2019) mentre la rete utilizzata per questo studio ha una densità molto maggiore. Questo dovrebbe permettere di raggiungere un livello di sicurezza nella validazione molto maggiore.

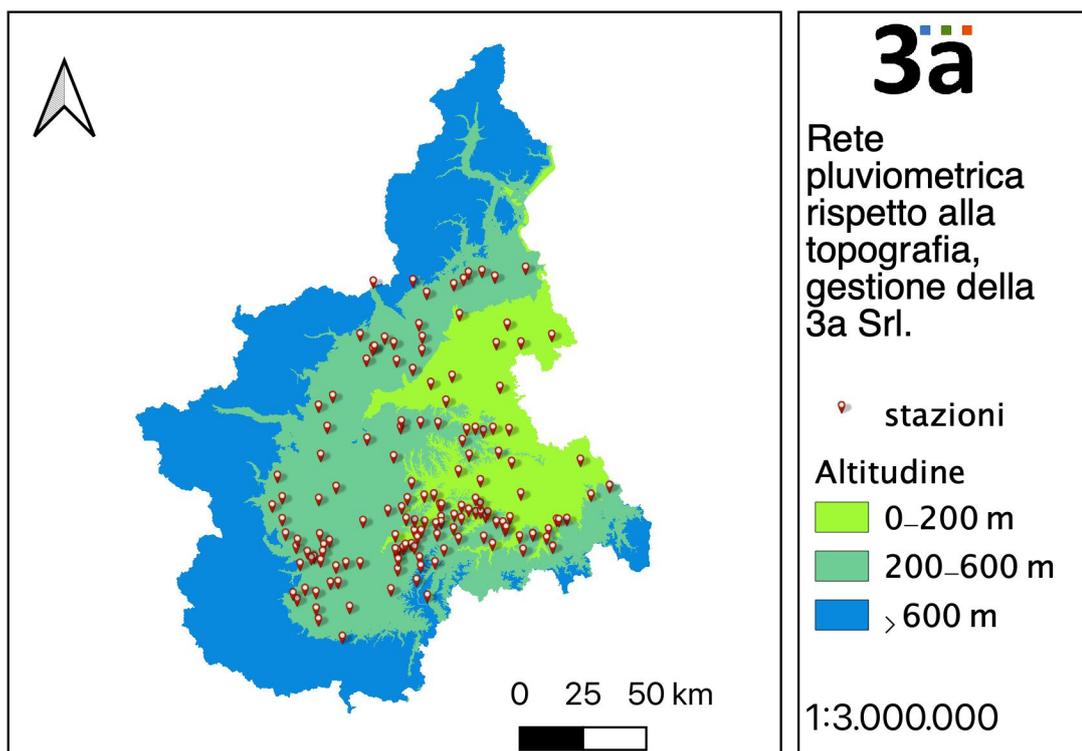


Figura 15 Posizioni delle centraline della rete pluviometrica, rappresentazione rispetto all'altitudine

Lo studio tuttavia, non ha come obiettivo la validazione del sistema GPM, quanto l'opportunità di sondarne l'applicabilità in situazioni che richiedano una precisione elevata, come ad esempio quella necessaria in ambito agricolo. Nella missione GPM non è specificato un target a cui si riferisce il servizio,

completamente gratuito, se non per la Final Run pensata per fini di ricerca. Il limitato tempo dopo il quale è possibile ottenere i primi risultati, lo rende intrinsecamente un meccanismo dalle forti potenzialità in applicazioni che necessitano di dati in continuo aggiornamento e in cui siano richieste risposte in tempi rapidi. Ad esempio, in campo agricolo la possibilità di avere valori di precipitazione ogni giorno potrebbe essere di supporto all'utilizzo di modelli previsionali su insetti e patologie della pianta. Potrebbe, inoltre, avere risvolti positivi nel computo di un bilancio idrico in continua evoluzione, aiutando gli agricoltori nella gestione irrigua e applicando una politica sostenibile di utilizzo della risorsa idrica. Sarebbe altresì possibile, grazie alla griglia con la quale sono distribuiti i dati elaborati da IMERG, sviluppare modelli previsionali ad hoc, sito-specifici e calibrati in base alle caratteristiche locali del suolo, alla vegetazione presente e alle condizioni di crescita delle colture ed ai meccanismi di deflusso delle acque, in modo da incrementare notevolmente l'utilizzo oculato delle risorse idriche e dei prodotti per il trattamento di patologie. Applicare queste moderne metodologie in zone distanti da centraline meteorologiche o in cui la distribuzione delle stesse risulti troppo rada, potrebbe rappresentare un'importante opportunità per lo sviluppo e la sostenibilità agricola. Per questo motivo si è scelto di utilizzare il prodotto Late Run di IMERG, in modo da poter contare sia sull'affidabilità di un'ulteriore analisi da parte dell'algoritmo, sia sui tempi ridotti a sole 12h per poter utilizzare i dati, con cui resta possibile produrre previsioni e applicare modelli con tempi adatti al settore agricolo: infatti, l'Early Run ha un tempo di processamento di cui ne risente parzialmente la qualità, mentre il Final Run, per quanto sia supposto essere più preciso, è pubblicato con un ritardo rispetto all'acquisizione non compatibile coi modelli agrometeorologici.

Una volta stabilite le necessità e la tipologia di dati da utilizzare, si è provveduto in seguito a raccogliere i dati per un periodo sufficiente.

La 3a possiede un database privato da cui è possibile ottenere i dati impostando una *query SQL*, ovvero seguendo il linguaggio informatico comunemente adottato per la ricerca delle informazioni in un database, per scaricare i dati filtrati via linea di comando in una tabella. Si è deciso, per semplicità, di ottenere i soli valori di pioggia, in quanto gli altri parametri non sarebbero stati in ogni caso confrontati con il GPM. I file "csv" sono organizzati in modo da avere dati relativi a stazione, orario di acquisizione e quantitativo di pioggia registrato. La localizzazione del punto di acquisizione dei dati, quindi, è registrata in un file separato logicamente: le regole di normalizzazione delle tabelle impongono di non ripetere, ove non necessario, una stessa informazione.

Per quanto riguarda GPM, invece, non essendo i dati contenuti in un database di semplice consultazione, è necessaria una procedura leggermente più strutturata. Nonostante il servizio sia pubblico, occorre effettuare un'iscrizione e saper trattare la tipologia di dati offerta. Una volta accettata la domanda di iscrizione al servizio,

i dati possono essere scaricati singolarmente o in blocco. Ognuno contiene informazioni riguardanti mezz'ora di acquisizione: il download di numerosi singoli file è un metodo ovviamente non consigliabile. È stato utilizzato quindi il software gratuito *Filezilla*, che permette il trasferimento di file in rete tramite il protocollo FTP. Dalla pagina principale di download dei file GPM <https://gpm.nasa.gov/data-access/downloads/gpm> (NASA 2020) si viene reindirizzati al server FTP

<ftp://jsimpson.pps.eosdis.nasa.gov/NRTPUB/imerq/early/> (NASA 2020), nella quale è possibile navigare tra le numerose directory: queste sono nominate secondo lo schema AAAAMM (anno-mese), al cui interno si trovano i file HDF contenenti la stima di precipitazione e diverse informazioni ausiliarie. Il formato HDF permette di mantenere queste informazioni in una struttura precisa, includendo eventuali metadati, e suddividendo in dataset relativi a specifiche categorie di dati o elaborazioni.

La stima vera e propria della precipitazione si trova nei dataset denominati *PrecipitationUncal* e in *PrecipitationCal*: la versione *cal* della versione Late Run ha subito la calibrazione tramite coefficienti climatologici variabili in base a mese e posizione, mentre nella Final Run la stima è tale per cui la somma delle misure 30-minutali eguaglia il totale mensile dell'unione delle misure satellite-pluviometri.

Per l'analisi in questa tesi è stata utilizzato il dataset *PrecipitationCal*. In Figura 16 qui sotto vi è una breve descrizione dei restanti dataset, estratta dalle note tecniche dell'algoritmo IMERG (Nasa Algorithm Theoretical Basis Document V06 2019).

<i>Half-hourly data file (Early, Late, Final)</i>	
<i>precipitationCal</i>	<i>Multi-satellite precipitation estimate with gauge calibration (recommended for general use)</i>
<i>precipitationUncal</i>	<i>Multi-satellite precipitation estimate</i>
<i>randomError</i>	<i>Random error for gauge-calibrated multi-satellite precipitation</i>
<i>HQprecipitation</i>	<i>Merged microwave-only precipitation estimate</i>
<i>HQprecipSource</i>	<i>Microwave satellite source identifier</i>
<i>HQobservationTime</i>	<i>Microwave satellite observation time</i>
<i>IRprecipitation</i>	<i>IR-only precipitation estimate</i>
<i>IRkalmanFilterWeight</i>	<i>Weighting of IR-only precipitation relative to the morphed merged microwave-only precipitation</i>
<i>probabilityLiquidPrecipitation</i>	<i>Probability of liquid precipitation phase</i>
<i>PrecipitationQualityIndex</i>	<i>Quality Index for precipitationCal field</i>
<i>Monthly data file (Final)</i>	
<i>precipitation</i>	<i>Merged satellite-gauge precipitation estimate (recommended for general use)</i>
<i>randomError</i>	<i>Random error for merged satellite-gauge precipitation</i>
<i>gaugeRelativeWeight</i>	<i>Weighting of gauge precipitation relative to the multi-satellite precipitation</i>
<i>probabilityLiquidPrecipitation</i>	<i>Accumulation-weighted probability of liquid precipitation phase</i>
<i>PrecipitationQualityIndex</i>	<i>Quality Index for precipitationCal field</i>

Figura 16 Elenco delle variabili e descrizione delle stesse per ogni prodotto in uscita da IMERG

3.2. Pre-processing

I dati provenienti dalle due fonti si presentano in modo molto differente: quelli della 3a in formato tabellare, mentre i dati IMERG in forma matriciale, ovvero disposti all'interno di una griglia con maglia $0,1^\circ$, corrispondente a queste latitudini ad un quadrato sulla superficie di poco maggiore ai 5km di lato. Inoltre, la risoluzione spaziale è differente: i primi rappresentano un accumulo orario, mentre i secondi un tasso di precipitazione semi-orario, ovvero 30 minutale.

Per poter confrontare i due mondi è necessario prima di tutto trasporne uno dei due nella stessa struttura dell'altro, per cui si è scelto di spazializzare sulla medesima griglia di $0,1^\circ$ i dati registrati dalle stazioni pluviometriche.

In Figura 17 è possibile osservare la distribuzione delle stazioni sul territorio piemontese. Si evidenzia in particolar modo come le stazioni si trovano principalmente in territorio pianeggiante, con poche eccezioni nella zona sud del Piemonte in cui è presente un ammasso collinare, e nella zona settentrionale della distribuzione, prossima ai rilievi alpini.

È stato necessario effettuare una serie di operazioni sequenziali prima di poter effettuare il confronto vero e proprio: si tratta del *pre-processing*. In questo caso esso consiste nel trasporre in forma matriciale i valori orari della rete pluviometrica, mentre i dati provenienti da IMERG si presentano già sotto forma di griglia, ovvero una matrice di numeri. L'algoritmo americano produce una stima di pioggia a livello mondiale, è stato quindi necessario servirsi di un software (Matlab) per estrarre le celle di interesse: prima è stata isolata un'area compresa tra le latitudini e le longitudini massime e minime in cui è racchiuso il Piemonte, successivamente è stata applicata una "maschera" in modo da ottenere all'esterno della regione valori "no data".

Per quanto riguarda le stazioni, queste occupano una posizione pseudo casuale sul territorio, con densità e distribuzione variabile, contrapposte ad una rigida griglia. È inevitabile che alcune stazioni si trovino all'interno di una stessa cella: questo causa una perdita di precisione del dato puntuale e costringe ad effettuare una media dei valori iscritti all'interno di una stessa cella. Questo metodo permette, tuttavia, di ottenere un valore rappresentativo di un'area più ampia e, per natura stessa del processo di mediazione, con estremi contenuti e andamento più morbido della serie temporale di dati.

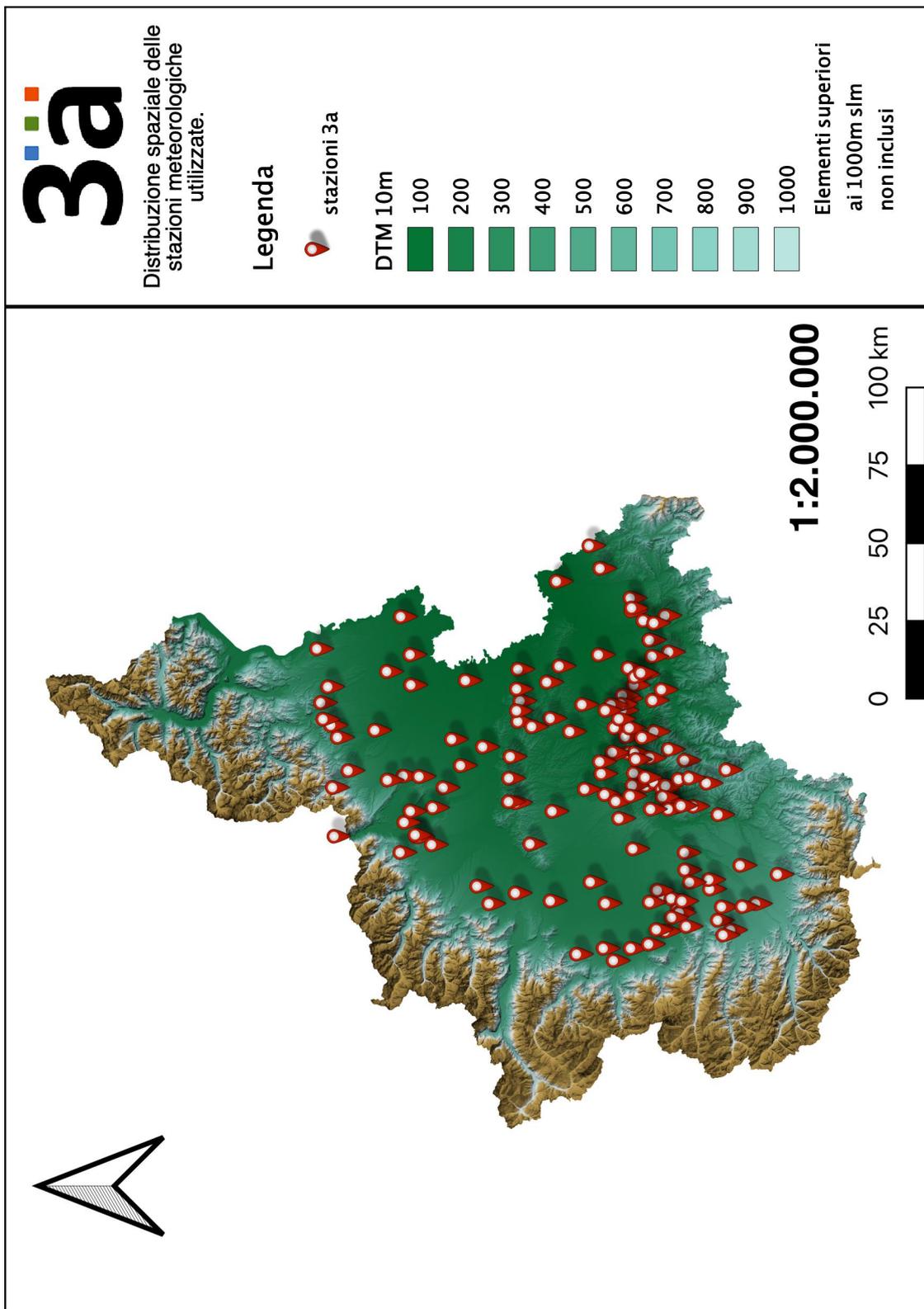


Figura 17 Rete Meteorologica in gestione alla 3a S.r.l.

Anche in questo caso è stato utilizzato Matlab per effettuare le operazioni di seguito riportate.

Alcune stazioni pluviometriche sono inserite all'interno di un programma per la prevenzione del gelo in campo ortofrutticolo, per cui registrano e inviano i dati ogni 10 minuti. Quindi è stato necessario calcolare l'accumulo orario di dette stazioni ed eliminare tutti i valori sub-orari.

In seguito, un passaggio importante è rappresentato dall'attribuzione di una stazione meteo in una cella piuttosto che in quella attigua: basandosi su studi precedenti si è scelto di utilizzare il metodo *nearest neighbor assignment*, ovvero l'assegnazione del dato alla cella il cui centro si trova in posizione più vicina alle coordinate della stazione (ESRI 2019). Nonostante si tratti della metodologia di assegnazione più semplice, è contemporaneamente la tecnica di ricampionamento più sicura, poiché non modifica il valore dei dati, nonché l'unica applicabile in questo caso in cui l'intento è di definire la posizione delle stazioni all'interno di una cella, in modo da poter confrontarne le misure nel tempo (in appendice è possibile consultare lo script utilizzato).

Questo metodo potrebbe potenzialmente essere utilizzato per eseguire una spazializzazione dei dati, ovvero effettuare un'assegnazione dei valori di pioggia senza che ne siano ricavati di nuovi tramite interpolazione, tuttavia non verrà discussa questa applicazione nella sezione. Successivamente è eseguito qualche piccolo passaggio intermedio di controllo e uniformazione dei vettori contenenti informazioni su data ed orario. Le celle rimaste vuote contengono valori "no data", rappresentati in Matlab come *NaN* (not a number), così da venire escluse da ogni successivo calcolo o elaborazione.

Il risultato di questi passaggi è visibile in Figura 18, dove in giallo vi sono le celle contenenti almeno una stazione e in blu il totale delle celle estratte da GPM, in seguito all'applicazione della maschera.

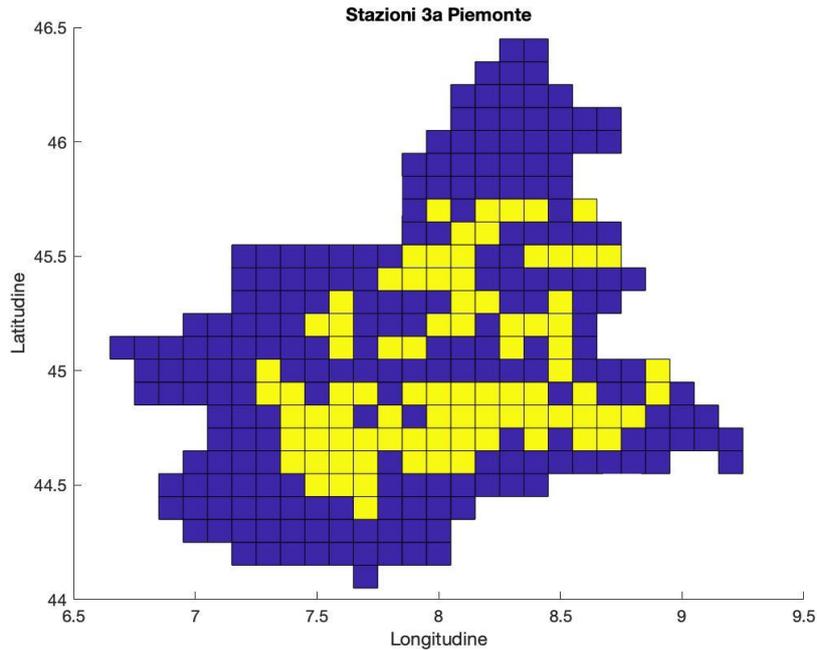


Figura 18 Posizione delle stazioni meteorologiche utilizzate disposte sulla stessa griglia di $0,1 \times 0,1^\circ$ in cui vengono elaborate le stime di IMERG. In giallo le caselle contenenti almeno 1 stazione, in blu le celle senza stazioni, non considerate nelle successive analisi

Entrambe le fonti si trovano a questo punto del pre-processing in forma matriciale: tuttavia sono presenti numerose matrici rappresentanti i dati di un istante specifico, un'ora in un singolo giorno. Per poter consultare in maniera agevole l'intera serie storica in tutte le celle si sono sovrapposte le matrici in modo tale da creare una matrice 3D, in cui, immaginando di trasporla in un piano cartesiano, nelle dimensioni x e y si estende geograficamente il territorio piemontese, mentre lungo la z troviamo il tempo. Il risultato finale è una matrice $25 \times 27 \times 8760$: infatti il Piemonte è interamente contenuto tra le longitudini $6,63^\circ$ (est) e $9,20^\circ$ (ovest) e le latitudini $44,06^\circ$ (sud) e $46,46^\circ$ (nord), ricordando che la dimensione della cella è di $0,1 \times 0,1^\circ$. Il valore 8760 invece rappresenta il numero di ore in un anno non bisestile (365×24). Grazie a questa disposizione, selezionando la casella, ovvero la combinazione di latitudine e longitudine conoscendone gli estremi, è possibile consultarne tutta la serie storica di un anno lungo la terza dimensione, così come osservare la distribuzione della precipitazione sul territorio in ciascuna ora.

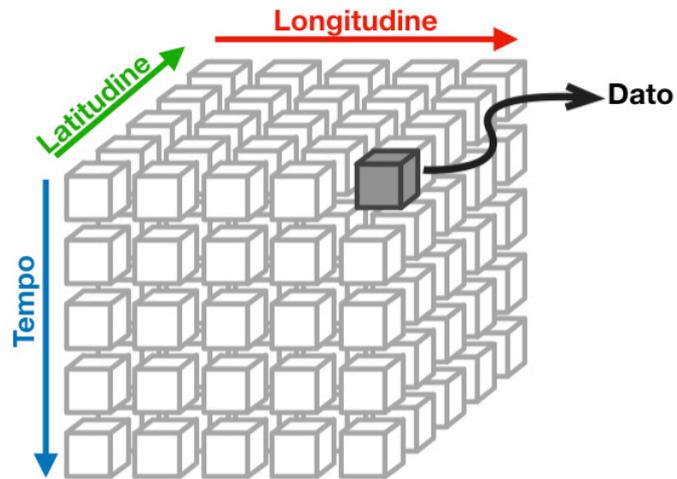


Figura 19 Rappresentazione schematica della costruzione della matrice di analisi. Si tratta di sovrapporre le mappe di distribuzione dei dati orari, in modo da poter consultare un particolare momento estraendo un piano orizzontale, oppure una serie storica in una posizione estraendo una colonna.

3.3. Processing

Per valutare la correttezza della stima della precipitazione operata dalla missione GPM e dall'algoritmo IMERG, l'operazione più semplice da effettuare è mettere a confronto le serie storiche: questo fornisce un'immediata idea dei principali punti di differenza e similitudine tra la stima e la realtà osservata e, eventualmente, se siano presenti picchi in una o nell'altra fonte che non trovano un'equivalente nell'altra. Inoltre, risulta semplice individuare la presenza di una costante sovra o sottostima. Per la costruzione di una serie di confronto è sufficiente selezionare una coppia di coordinate, ovvero la casella corrispondente e, grazie alla disposizione adottata per spazializzare i dati, estrarre la dimensione z dalla matrice tridimensionale: si avrà in tal modo la successione di dati riguardanti una stessa posizione nel tempo.

Per questo metodo e per i seguenti si sono usate le seguenti aggregazioni temporali:

- Oraria
- Giornaliera (24 ore)
- Tri-giornaliera (72 ore)
- Mensile

Il passo immediatamente successivo è la costruzione di uno scatter plot per valutare la correlazione tra i due set di misure: ovviamente nel caso di identità o lieve scostamento la nuvola di punti dello scatter sarà allungata lungo la retta di riferimento di equazione $y=x$.

Infine, sono stati calcolati diversi indici di natura statistica e categorica. Tra i coefficienti statistici vi sono il *root mean square error* o RMSE:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i^p - p_i^s)^2}{n}}$$

dove

p^p : è l'i-esimo valore di pioggia registrato da pluviometro,

p^s : è l'i-esimo valore di pioggia registrato da satellite,

n: numero di elementi considerati;

e l'indice di correlazione di Pearson:

$$r = \frac{\sum_{i=1}^n (p_i^p - \bar{p}^p)(p_i^s - \bar{p}^s)}{\sqrt{\sum_{i=1}^n (p_i^p - \bar{p}^p)^2} * \sqrt{\sum_{i=1}^n (p_i^s - \bar{p}^s)^2}}$$

Il RMSE rappresenta un modo per valutare quanto si discostino le stime di precipitazione dal valore reale di precipitazione, che supponiamo essere quello misurato dalla rete pluviometrica. Per ogni valore i -esimo, ovvero per ogni valore orario, si calcola la differenza (o scarto) tra le due misure, poi si esegue una media sommando i singoli scarti orari e dividendo per il numero totale di i (ovvero le ore in un anno, 8760) e si ricava la radice. Si tratta quindi della deviazione standard dell'errore di previsione o stima.

L'indice di correlazione di Pearson è il coefficiente più comune per misurare la relazione incrementale tra due misure, o due set di dati come in questo caso. Un coefficiente pari a 1 indica la massima correlazione, cioè per ogni incremento positivo di una variabile vi è un incremento positivo dell'altra variabile. Al contrario una correlazione pari a 0 equivale a dire che i dataset considerati non sono correlati. È anche possibile che la correlazione sia negativa: in tal caso ad un aumento della variabile A corrisponde una riduzione della variabile B.

Per valutare la performance e l'accuratezza di IMERG sono stati utilizzati anche alcuni indici categorici: POD, FAR, CSI e BIAS. Questi indici sono utilizzati dal gruppo NASA-JAXA stesso nella fase di validazione dell'algoritmo (Nasa Algorithm Theoretical Basis Document V06 2019).

Per calcolare gli indici è necessario considerare le categorie prestazionali, ricavate tramite la seguente tabella di contingenza (Mazzoglio 2019):

Tabella 3 Tabella di contingenza provvista di soglia, per identificare all'interno di categorie prestazionali ogni stima oraria. La soglia utilizzata è pari a 0,2mm per dati orari e aggregazione giornaliera e tri-giornaliera, 2mm per l'aggregazione mensile. Sotto tale soglia la stima viene considerata come 0. In verde i dati della rete pluviometrica, indicati anche come Pp (pluvio), in blu i dati GPM, o Ps (satellite). Di conseguenza se $Pp < \text{soglia}$ non è stato osservato alcun evento, se $Ps < \text{soglia}$ non è stato stimato alcun evento, ecc. Se un evento stimato è stato anche osservato, allora si tratta di un "hit", una misura corretta. Con lo stesso meccanismo troviamo le categorie "correct negatives" (evento non stimato e non osservato), "misses" (evento osservato ma non stimato), "false alarms" (evento stimato ma non osservato).

	Pp < soglia	Pp > soglia	Pluviometri
Ps < soglia	Correct negatives	Misses	Estimated Non events
Ps > soglia	False Alarms	Hits	Estimated Events
IMERG	Observed Non-events	Observed events	

La tabella valuta il rapporto tra stima e dato esatto: la stima proviene dall'algoritmo di valutazione dei dati satellitari, mentre l'esattezza è rappresentata dalla rete pluviometrica. Di conseguenza, impostata una soglia arbitraria variabile in base all'aggregazione temporale, è possibile individuare le seguenti categorie:

- *Correct Negatives*: non è misurata né stimata alcuna precipitazione. È il caso più comune nei dati orari alle latitudini piemontesi.
- *Misses*: è presente pioggia, ma la stima non ne rileva. È a tutti gli effetti una mancata stima di precipitazione.
- *False Alarms*: viene stimato un quantitativo di pioggia ma non viene misurata precipitazione a terra.
- *Hits*: viene stimato un quantitativo di pioggia e la precipitazione è reale.

L'utilizzo di una soglia è necessario per superare il problema della risoluzione differente nei due sistemi. Inoltre, permette di includere tra i correct negatives quelle misure prossime allo 0 di cui non è certo il fenomeno piovoso.

Le soglie usate sono:

- 0,2 mm per orari, aggregazione giornaliera e tri-giornaliera
- 2 mm per i mensili

Nonostante il valore di correct negative sia il più comune, è il meno indicativo. La situazione ideale è una massimizzazione degli hits e un valore prossimo a 0 di falsi e mancati allarmi, i quali sono parimenti negativi. Infatti, sia nel caso di falso che mancato allarme la stima è certamente errata. Tuttavia, per la natura della categorizzazione, la stima può essere errata anche nel caso di un hit: infatti il quantitativo di pioggia misurata non è considerato in questo caso, si tratta infatti di una questione di rilevazione del fenomeno, invero una valutazione più grezza ma di grande efficacia per valutare la performance.

Si riportano due esempi per spiegare esaurientemente la necessità di considerare due tipologie di indici. Si immagini di dover valutare una singola cella della spazializzazione di IMERG: è possibile, ad esempio, che vi sia un basso RMSE, ma che il satellite abbia stimato la presenza di poca pioggia tutti giorni, ma guardando alla precipitazione media tra giorni asciutti e piovosi, la stima totale non si discosti troppo da quella reale. Tuttavia, ha effettivamente piovuto in un numero limitato di giorni, in quantità ben superiori alla stima giornaliera media dell'algoritmo. In un simile caso l'utilizzo di indici categorici risulterebbe fondamentale. Oppure, si prenda in considerazione un periodo casuale di 2 mesi molto poco piovosi: immaginiamo che il satellite stimi sempre 0mm di pioggia, ma in un giorno del periodo ci sia stato un singolo fenomeno piovoso di media entità (es. 10-20 mm/giorno) Se si guardasse soltanto le differenze di misura potremmo vedere qualche millimetro di differenza (sull'intero periodo) e pensare che la stima non sia esageratamente sbagliata, risultando solo 10mm di differenza su due mesi, ma in realtà il sistema di stima non è stato in grado di rilevare l'evento piovoso, seppur modesto. Per questo motivo gli indici categorici sono da utilizzare congiuntamente agli indici statistici.

Tra gli indici computati si trova il POD, o *Probability Of Detection*, che esprime il rapporto tra gli eventi piovosi rilevati come tali e il totale di eventi di pioggia. In

altre parole, quante volte il sistema si è accorto della precipitazione sulla cella. Il valore ottimale è 1.

$$POD = \frac{hits}{hits + misses}$$

Da notare l'utilizzo al denominatore di *hits+misses*: la somma indica esattamente il numero di casi di precipitazione osservata dalla rete pluviometrica, ad esclusione degli 0 (o valori sotto soglia).

Il FAR, ovvero *False Alarm Ratio*, rappresenta una stima dell'affidabilità del sistema di stima tramite il rapporto fra i falsi eventi osservati e il numero totale di eventi stimati, rispondendo alla domanda: quanti di essi sono falsi? Di conseguenza è auspicabile che tenda a 0.

$$FAR = \frac{false\ alarms}{hits + false\ alarms}$$

Il *Critical Success Index*, abbreviato in CSI, è il numero di eventi stimati correttamente rapportato al totale degli eventi stimati: quindi se pari ad 1 gli eventi stimati in modo esatto ne rappresentano il totale.

$$CSI = \frac{hits}{hits + misses + false\ alarms}$$

L'ultimo parametro, detto hit BIAS, è una variazione del normale BIAS presentandosi in versione frazionaria, ovvero si tratta dello scostamento tra il numero di eventi di precipitazione stimati e il numero di eventi osservati, ovvero

$$hitBIAS = \frac{hits + false\ alarms}{hits + misses}$$

Se pari ad 1, significa che gli eventi stimati sono pari in numero agli eventi misurati.

Tutte le analisi sono state eseguite utilizzando il prodotto Late Run di IMERG.

4. Risultati

Sono stati analizzati in dettaglio i dati relativi al periodo compreso tra il 2013 e il 2017. In seguito al reperimento degli stessi, il pre-processing ha permesso di porre su uno stesso piano di confronto valori di precipitazione originariamente non paragonabili per la natura dell'acquisizione, ovvero è stata effettuata una spazializzazione dei dati registrati dalla rete pluviometrica con la stessa maglia di $0,1 \times 0,1^\circ$ adottata da GPM. Il confronto si è limitato alle aree in cui entrambi gli strumenti utilizzati hanno registrato dati, escludendo quelle aree con serie storiche incomplete o in cui uno dei due strumenti non è presente, permettendo di ottenere una migliore accuratezza. I risultati saranno esposti a partire dagli indici categorici, di più immediata interpretazione.

Al termine dell'analisi prestazionale è emersa chiaramente la relazione tra l'aumento del tempo di aggregazione (da orario a mensile) ed il miglioramento sia degli indici statistici sia di quelli categorici. Guardando al 2016 ad esempio, in Figura 20, è evidente il netto cambio nella capacità di rilevare la precipitazione da parte del GPM tra valori orari e giornalieri: nei primi, infatti, le stime corrette sono inferiori a quelle in cui la pioggia non è stata rilevata e il numero dei falsi allarmi è elevato. Guardando all'aggregazione giornaliera la situazione risulta migliore, le stime corrette crescono notevolmente, nonostante rimanga alto il numero dei falsi allarmi e dei fenomeni mancati. Occorre precisare che è stata utilizzata la tabella di aggregazione in Tabella 3, provvista di una soglia limite per evitare che siano considerati come hit anche i valori molto prossimi allo 0 e al di sotto della sensibilità dei pluviometri.

È possibile considerare positivo il risultato della stima dall'aggregazione temporale tri-giornaliera in avanti, quando gli hits superano la somma di misses e falsi allarmi, giungendo infine alla visualizzazione mensile che, considerando soltanto 12 valori (12 mesi) per cella, presenta una quasi totalità delle rilevazioni corrette, complice del fatto che sia sufficiente rilevare un solo evento piovoso al mese per rendere il mese un *hit*. È importante ricordare la natura degli indici categorici: il 100% di eventi piovosi rilevati non significa che siano stati rilevati correttamente, ma soltanto che siano stati rilevati, nonostante rappresenti un motivo per valutare positivamente la prestazione.

Possiamo inoltre evidenziare i particolari andamenti dei parametri. Ad esempio, i *correct negatives* in blu posseggono un trend decrescente all'aumentare dell'aggregazione: questo fenomeno è causato dalla scarsa probabilità di osservare serie temporali estese con assenza di pioggia. È sufficiente che sia rilevato un singolo evento di precipitazione all'interno del periodo considerato perché l'intero periodo sia escluso dai *negatives*. Seguendo il medesimo meccanismo, si può osservare il trend crescente degli hits, poiché rilevare correttamente un fenomeno piovoso all'interno del periodo di aggregazione rende l'intero periodo corretto.

Si osserva invece un'assenza di monotonia nell'andamento di misses e false alarms: possiamo attribuire la loro bassa percentuale nei valori orari all'enorme mole di valori attribuiti ai correct negatives, per cui essi risultano in netta minoranza, mentre nell'aggregazione mensile si può ricondurre la loro scarsa presenza alla stessa spiegazione fornita per gli hits: un solo evento rilevato correttamente esclude automaticamente misses e false alarms. Andamenti simili sono osservabili in tutti gli anni considerati nell'analisi, consultabili in appendice. Nella parte inferiore della Figura 20 si può apprezzare l'evoluzione degli indici realizzati sulla base delle prestazioni valutate nella tabella di aggregazione. Il FAR indica l'occorrenza di falsi allarmi sugli eventi osservati, per cui il trend decrescente rappresenta la capacità del sistema di generare un numero inferiore di finti eventi piovosi all'aumentare dell'aggregazione. Il POD invece, valuta la probabilità di rilevare l'evento, rappresentando l'indice di prestazione più immediato e di maggior importanza. Si osserva, nei dati orari, come l'indice sia al di sotto del 50%, mentre dall'aggregazione giornaliera in avanti risulta maggiormente probabile che il sistema satellitare sia in grado di rilevare l'evento. Per comprendere quanti degli eventi rilevati trovino un riscontro nella sorgente di confronto occorre osservare il CSI, il quale paragona gli hits al numero di eventi rilevati dall'algoritmo. Il suo valore ideale è l'unità, indice del fatto che tra tutti gli eventi rilevati non vi sono stati né falsi allarmi né eventi mancati. Nei valori orari in particolare, è possibile notare con semplicità come la percentuale di hits, false alarms e misses sia molto simile, producendo un CSI non superiore al 30%. Osserviamo infine il BIAS, calcolato come rapporto tra gli eventi stimati e gli eventi reali, il quale indica la capacità del sistema di stimare correttamente la presenza di pioggia, nel momento in cui vi sia una stima effettiva. Tutti gli indici, ad esclusione del FAR, presentano un'evoluzione crescente all'aumentare dell'aggregazione temporale, raggiungendo valori prossimi ad 1 per l'intervallo mensile.

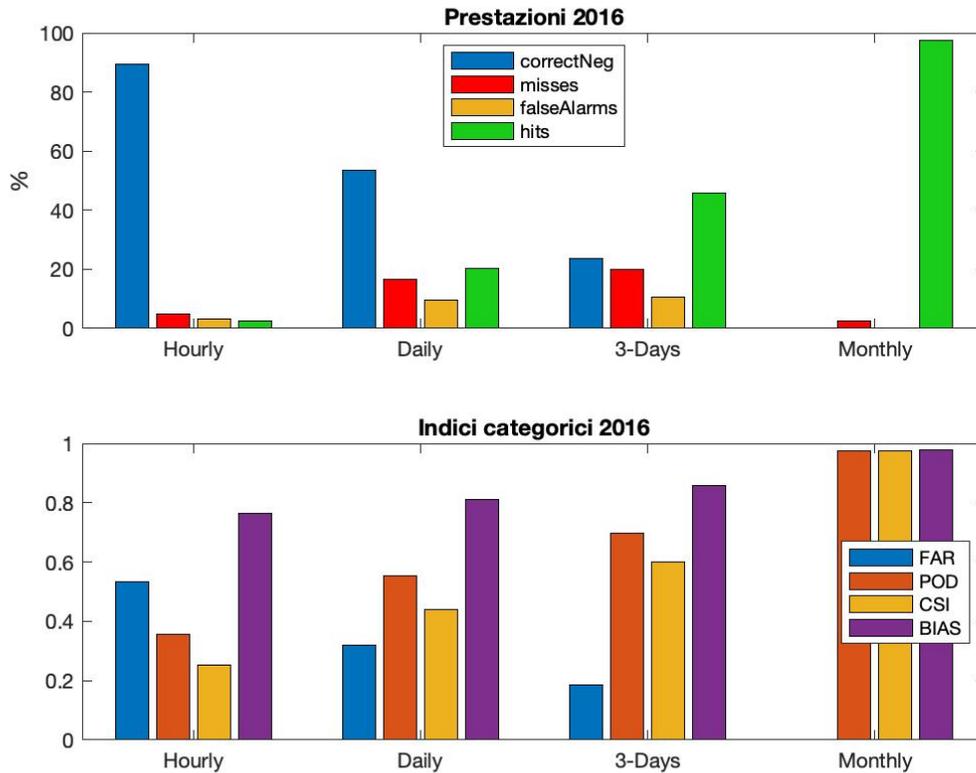


Figura 20 Risultati della tabella di contingenza (Tabella 3): all'aumentare dell'aggregazione temporale sull'asse delle ascisse diminuisce la probabilità di incorrere in periodi senza precipitazione, per cui i correct negatives subiscono una riduzione; contemporaneamente gli hits aumentano, poiché è sufficiente rilevare una singola precipitazione all'interno del periodo di aggregazione per classificare come corretta la stima. Al di sotto invece si trova l'evoluzione degli indici categorici al variare del periodo di aggregazione: il valore ottimale è pari a 0 per il FAR e pari ad 1 per gli altri indici, evidenziando un comportamento positivo della stima.

Si riportano in appendice gli stessi schemi prodotti per i differenti anni analizzati. È possibile notare come nel 2015 (Figura 43) e nel 2017 (Figura 44) le prestazioni siano leggermente inferiori: tuttavia è possibile si tratti di una semplice flessione statistica.

È stato preso in considerazione anche l'output della Final Run di IMERG per l'anno 2016, ovvero il prodotto dell'algoritmo disponibile al pubblico dopo un lasso di tempo di 3 mesi e mezzo, necessario per eseguire una serie di correzioni sulla stima di pioggia basate sulla stagione, la posizione geografica ed i dati raccolti da una rete pluviometrica a terra. Il confronto tra i risultati della Late Run e quelli della Final Run è necessario per indagare la presenza di sostanziali differenze e l'influenza della correzione attuata grazie alla rete a terra. Confrontando la Figura 21 con la Figura 20, identiche nei contenuti e nei comportamenti degli indici ma riferite rispettivamente alla Final e alla Late run, si possono evidenziare lievi miglioramenti generali in tutti gli indici, tuttavia essi non si ritengono tali da giustificare il tempo di oltre 3 mesi necessario per produrre la versione Final, rispetto alle 12 ore occorrenti alla versione Late. La qualità della Late Run è stata

considerata sufficientemente precisa in relazione alle versioni rese disponibili da IMERG, per cui le seguenti analisi si riferiranno esclusivamente a questa versione.

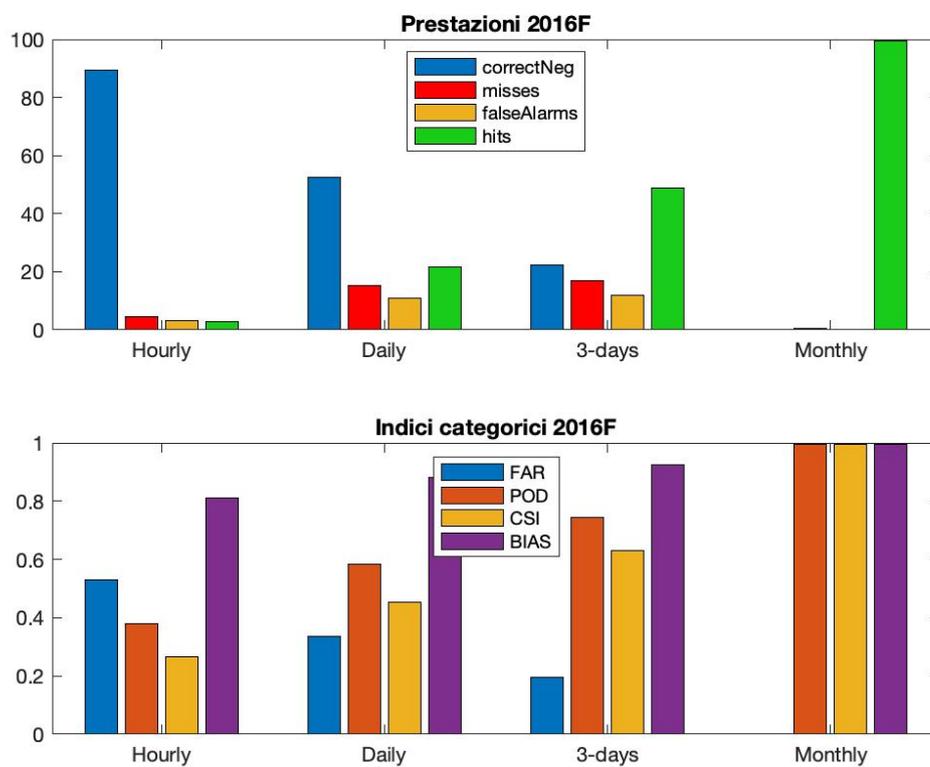


Figura 21 Risultati della tabella di contingenza e degli indici categorici, entrambi elaborati a partire dai dati della Final Run, il prodotto finale e di maggior accuratezza fornito dalla reiterazione dell'algoritmo IMERG e dalla correzione della stima utilizzando dati di una rete pluviometrica, 3 mesi e mezzo dopo l'acquisizione dei dati. Rispetto agli stessi indici elaborati grazie ai dati della Late Run (12 ore dopo l'acquisizione, senza interazione con rete pluviometrica) e visibili in Figura 20, vi è un miglioramento appena apprezzabile.

Gli indici categorici rappresentano una valutazione della capacità del sistema GPM di rilevare un evento piovoso, con quale probabilità e quando vi è un riscontro con una precipitazione reale, ovvero i valori di pioggia registrati della rete pluviometrica sono anch'essi sopra alla soglia di rilevamento. Tuttavia, per loro natura questi indici, una volta appurato che sia il sistema in remoto sia quello sulla superficie hanno registrato un evento piovoso, non sono in grado di quantificare l'accuratezza della misura, ovvero il discostamento in mm di pioggia.

È necessario quindi osservare nella prossima sezione la distribuzione della radice dello scarto quadratico medio o RMSE: questo indice rappresenta una misura dell'accuratezza della stima satellitare, esprimendo lo scarto medio tra il valore corretto e quello stimato. Valori di RMSE prossimi allo zero sono indicatori di una stima accurata. È necessario, tuttavia, considerare il suo valore in relazione al fenomeno a cui si riferisce, ad esempio valutando la precipitazione media rilevata

dal satellite e dalla rete pluviometrica per ogni cella di $0,1^\circ \times 0,1^\circ$. In Tabella 4, è possibile apprezzare la vicinanza tra le due misure.

Tabella 4 Precipitazione media rilevata in ogni cella di lato $0,1^\circ \times 0,1^\circ$

	Orari	Giornalieri	3-Giorni	Mensile	
Pluviom.	1,40	6,96	11,67	77,03	Media
Satellite	1,51	2,64	12,12	80,90	
Pluviom.	59,40	255,60	385,20	423,60	Max
Satellite	84,16	327,86	399,88	497,41	

In Figura 22, si trova un boxplot di Matlab, uno strumento utile per visualizzare sinteticamente le statistiche di un parametro: il tratto rosso al centro del rettangolo indica la misura mediana; i limiti del rettangolo rappresentano il 25° e il 75° percentile della distribuzione; i tratti orizzontali esterni sono i valori rispettivamente di massimo e minimo, mentre il “+” rosso indica i valori considerati “outliers”, ovvero quei singoli valori esterni alla distribuzione. Da questo tipo di grafico si evince che lo scarto medio è minimo per i valori orari, appena superiore allo 0. Lo scarto rimane accettabile per valori giornalieri, dove la media, una volta esclusi i valori sotto soglia, è di 6,95mm.

In Figura 23, in cui ogni colonna rappresenta una cella, si nota come i picchi in ogni aggregazione stazionino sempre sulle stesse celle, o per lo meno le celle con un alto RMSE nei valori orari rimangono alle prime posizioni nelle aggregazioni successive. Questo risulta chiaro dalle mappe in Figura 25, le quali indicano il valore di RMSE spazializzato nel 2016. Non è più visibile l’intera distribuzione delle celle GPM poiché sono state considerate le sole celle contenenti almeno una stazione.

Risultati

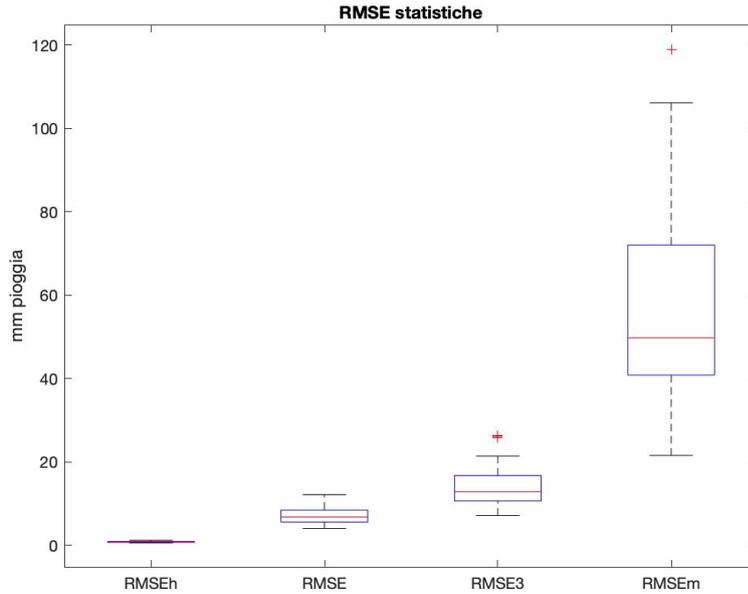


Figura 22 : Boxplot dell'errore quadratico medio, o RMSE, elaborato per differenti aggregazioni temporali per l'anno 2016. Il tratto rosso al centro indica la mediana, mentre i limiti del rettangolo rappresentano il 25° e il 75° percentile della distribuzione, i tratti orizzontali esterni sono i valori massimo e minimo, mentre il + rosso indica i valori considerati "outliers". Considerando periodi di aggregazione maggiori, la differenza tra valori massimi e minimi aumenta, indicando la scarsa accuratezza della stima di IMERG nel rilevare il quantitativo di precipitazione corretto

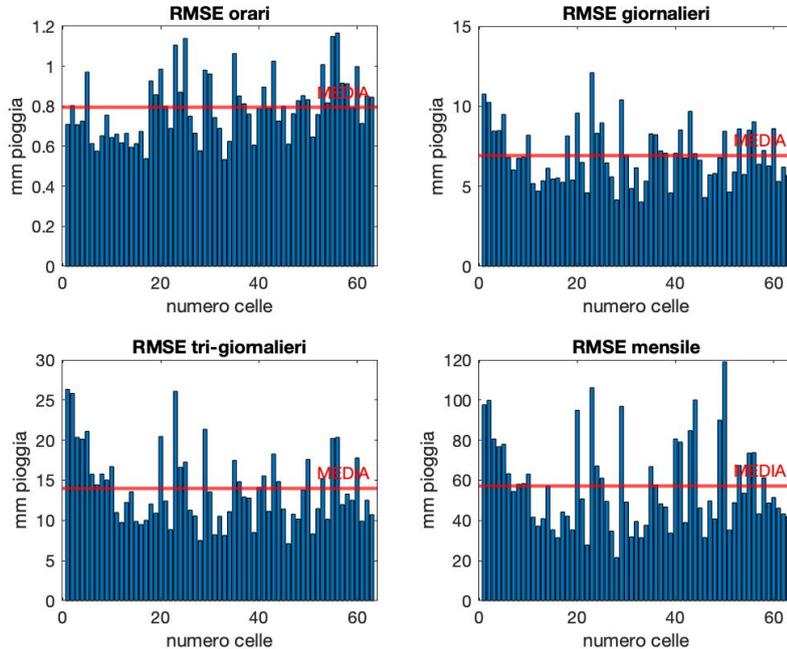


Figura 23 Distribuzione dei valori di RMSE per l'anno 2016, divisi per aggregazione temporale. Ogni colonna rappresenta lo scarto quadratico medio in mm di pioggia per una cella di area $0,1 \times 0,1^\circ$, il cui numero è sull'asse delle ascisse. In rosso è visibile la media. All'aumentare dell'aggregazione troviamo scarti maggiori, tuttavia occorre osservare come alcune delle celle con un elevato RMSE nei valori orari siano le stesse che spiccano nell'aggregazione mensile.

Avendo una più definita idea dei limiti inferiore e superiore dei valori di pioggia registrati, è possibile valutare correttamente il significato e l'entità dello RMSE. Esso risulta, infatti, paragonabile alla media stessa. Osservando però i valori massimi di precipitazione in ogni aggregazione, e immaginando che la distribuzione sia di conseguenza da 0 al valore massimo, lo scarto quadratico medio permette di asserire che la capacità di stimare correttamente la quantità di pioggia è limitata, ma potrebbe risultare accettabile per altre applicazioni che esulano dalle applicazioni di precisione. Infatti, in media la stima satellitare e la misura pluviometrica risultano simili, suggerendo la possibilità di sfruttare i prodotti di IMERG per spazializzazioni regionali con intervalli di tempo più ampi, quali ad esempio aggregazioni mensili o stagionali.

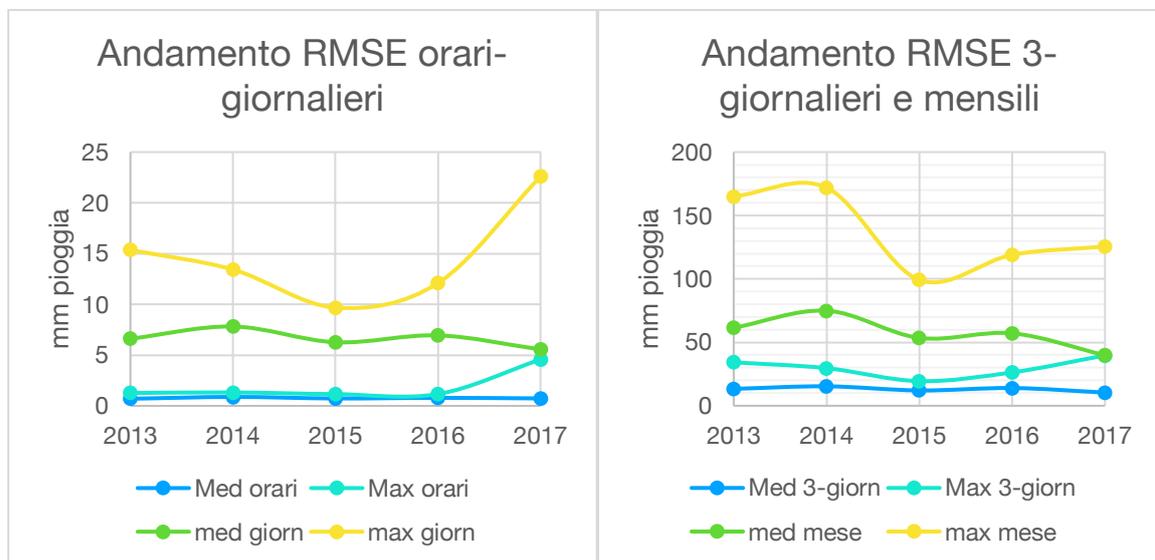


Figura 24 Andamento RMSE a confronto

Un altro parametro fondamentale per lo studio è la distribuzione spaziale dell'errore, rappresentata in Figura 25.

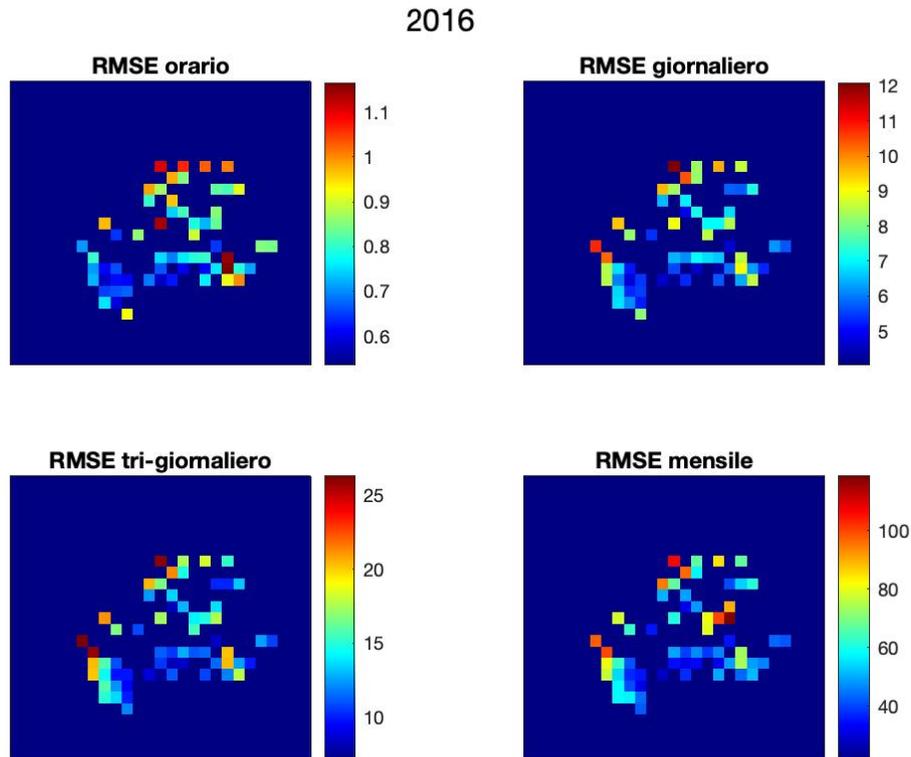


Figura 25 Distribuzione spaziale dei valori di RMSE per l'anno 2016: i valori aumentano dal blu al rosso. I valori maggiori di scarto si trovano nella zona a nord-ovest, in prossimità dell'arco alpino, e a sud-est dove le stazioni incluse nella cella hanno un'altitudine maggiore. Al contrario, al centro della distribuzione, dove è prevalente un'area pianeggiante, i valori di scarto sono nettamente inferiori.

Si può osservare una prestazione peggiore nelle celle site a nord della distribuzione, con un apprezzabile distacco rispetto a quelle nella zona del Piemonte centrale, ad esempio. Ricordando la mappa in Figura 17, occorre sottolineare che le stazioni più settentrionali si trovano nella zona pedemontana a nord di Torino, in provincia di Biella.

Guardando alla Tabella 5, possiamo osservare come in tutti gli anni vi sia un RMSE medio giornaliero di tutte le celle compreso tra 5,5 e 7,8 mm di pioggia. Questo significa che in media, la stima si discosta di 6,64 mm dall'esattezza.

Tabella 5 Valori medi e massimi di RMSE nelle differenti aggregazioni temporali

	Orari	Giornalieri	Tri-giornalieri	Mensili	
Media	0,704	6,616	13,322	61,452	2013
Max	1,268	15,359	34,256	165,086	
Media	0,883	7,820	15,334	74,942	2014
Max	1,324	13,428	29,659	171,932	
Media	0,733	6,270	12,082	53,478	2015
Max	1,158	9,680	19,416	99,150	
Media	0,794	6,946	13,978	57,138	2016
Max	1,166	12,098	26,331	119,004	
Media	0,736	5,566	10,431	39,817	2017
Max	4,571	22,593	39,654	125,588	

In appendice è possibile osservare l'evoluzione negli anni in ordine cronologico. Rispetto alla mappa del 2016 proposta in Figura 25, in cui la scala cromatica è stata adattata tra valore minimo e massimo, in quelle poste in appendice i limiti della scala sono gli stessi in tutte le rappresentazioni per semplificare l'interpretazione. È possibile confermare che i valori maggiori di RMSE si presentano nella zona pedemontana a nord in ciascun anno considerato nell'analisi, suggerendo che vi sia una relazione di causalità tra l'elevazione del terreno e la precisione della stima pluviometrica. In Figura 26 è visibile uno zoom della locazione delle celle la cui stima presenta oscillazioni maggiori.

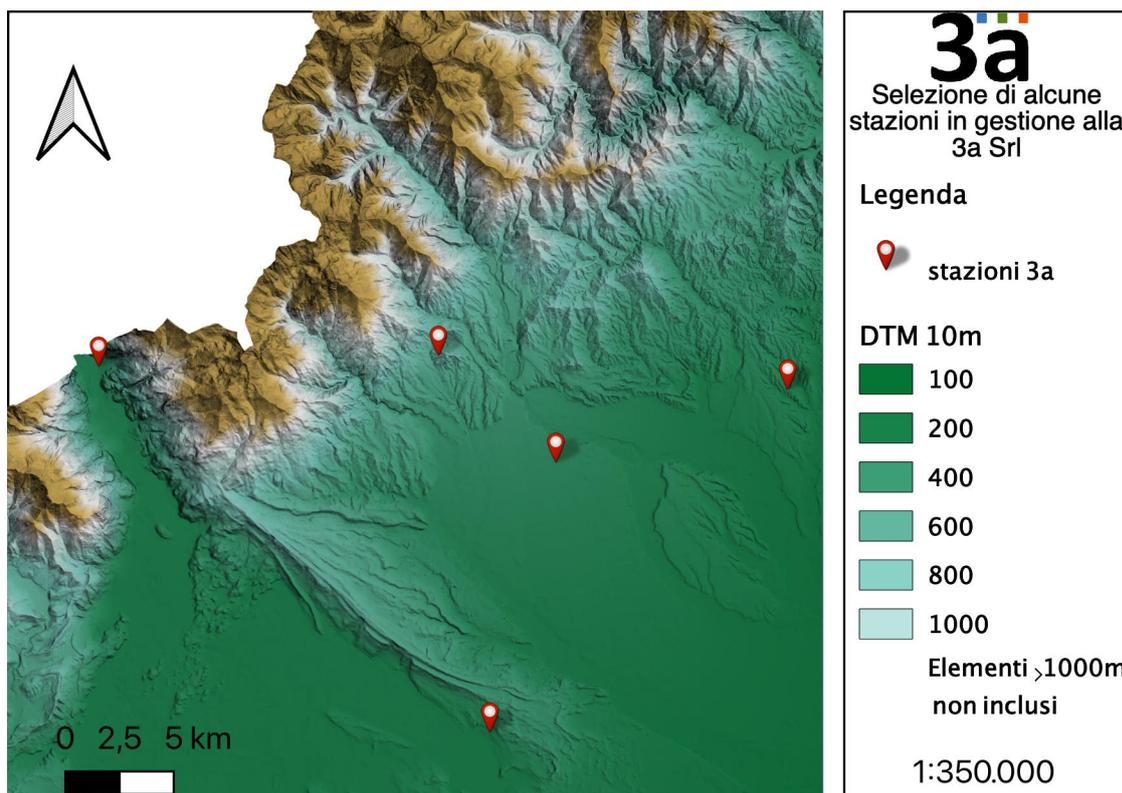


Figura 26 Visuale ingrandita di alcune stazioni poste a nord: le celle di $0,1^\circ$ di lato in cui sono state incluse durante la spazializzazione presentano RMSE maggiori e differenze più elevate tra stima e misura della centralina pluviometrica, evidenziando la difficoltà del sistema GPM di stimare correttamente la precipitazione in presenza di aree collinari o di rilievi montuosi.

Un'altra zona in cui si trova un RMSE in generale più elevato risulta essere a sud-est della distribuzione: si tratta nuovamente di una zona collinare, in cui la quota delle stazioni è stata riportata in Figura 27. La distribuzione delle stazioni evidenziate, comparata alla distribuzione dell'intera rete pluviometrica, mostra come la quasi totalità delle stazioni a quota superiore ai 600 metri si trovi nell'area in cui l'accuratezza della stima subisce inflessioni negative. Occorre ricordare che per poter operare il confronto le stazioni sono state ridistribuite in una rete di $0,1 \times 0,1^\circ$, per cui il valore di pioggia ricavato dalla rete pluviometrica è il frutto di una media di tutte le stazioni all'interno della cella, per cui l'effetto della quota è smorzato. Purtroppo, per questo stesso motivo, non è possibile individuare una precisa altitudine con cui operare e costruire una funzione di correlazione tra errore medio e quota, risulta tuttavia inopinabile l'effetto dell'altitudine sulla qualità della stima GPM.

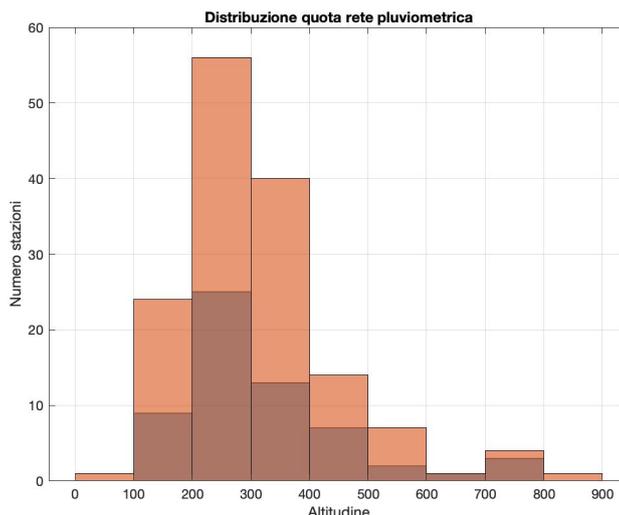


Figura 27 Distribuzione dell'altitudine delle stazioni poste nell'area collinare a sud del Piemonte (colore più scuro) rispetto all'intera rete pluviometrica (colore chiaro)

Per la valutazione è stato ricavato successivamente il coefficiente di Pearson (indicato talvolta con la lettera r minuscola): questo indice ci permette di comprendere la relazione tra i trend di due misure. Infatti, un coefficiente di Pearson positivo indica che ad un incremento di A corrisponde una crescita proporzionale del fattore B, mentre un coefficiente negativo rappresenta una riduzione di B rispetto ad un incremento di A.

L'aumento del coefficiente, in questo caso, corrisponde alla rilevazione di un quantitativo positivo di pioggia, per cui l'indice sarà sempre positivo.

Nelle rappresentazioni è stata mantenuta una scala fissa tra 0 e 1 per poter confrontare tra loro i differenti anni analizzati; si è scelto tuttavia di riportare nel testo principale il solo anno 2016 in quanto ritenuto maggiormente rappresentativo, in Figura 52. È possibile apprezzare un netto aumento generale in tutta l'area piemontese del valore dell'indice di Pearson considerando aggregazioni temporali maggiori. Osserviamo come i dati orari abbiano una scarsa correlazione, per cui l'indice risulta inferiore a 0,5 nella quasi totalità delle celle della griglia, fenomeno probabilmente dovuto al fatto che tra stima e misura bastano pochi minuti di differenza o un leggero shift nella registrazione del valore per cui in una determinata ora le misure non corrispondano. Passando all'aggregazione giornaliera è facilmente osservabile un generale incremento dei valori dell'indice: aumentando la finestra temporale diminuisce la probabilità che le quantità d'acqua cadute al suolo vengano considerate in un giorno adiacente, ma il coefficiente rimane discretamente ridotto, oscillando tra 0,4 e 0,6, presentando picchi maggiori in casi isolati. Nelle successive aggregazioni di 3 giorni e mensili non sono rari indici superiori allo 0,8, valore molto positivo.

Sono tuttavia presenti alcune celle in cui non è apprezzabile alcun deciso miglioramento dell'indice. Questo fenomeno è da attribuire all'aggregazione

stessa: sommando molteplici piccole differenze per costruire, ad esempio, l'aggregazione mensile, è plausibile che nel risultato la misura pluviometrica e la stima satellitare differiscano notevolmente.

Riguardo alla distribuzione spaziale, a differenza dell'indice di RMSE, in questo caso non è di eguale immediatezza individuare un blocco di celle che mantengono un andamento positivo o negativo nel tempo. Tuttavia, è possibile effettuare un paragone con le celle nell'area pedemontana nella zona a nord della rappresentazione e nell'area collinare posta a sud-est: anche nel caso dell'indice di Pearson queste aree presentano valori inferiori, indicando una minore accuratezza della stima satellitare. Questo fenomeno è osservabile in tutti gli anni considerati, ad esclusione del 2017 in cui l'indice ha una distribuzione pseudocasuale.

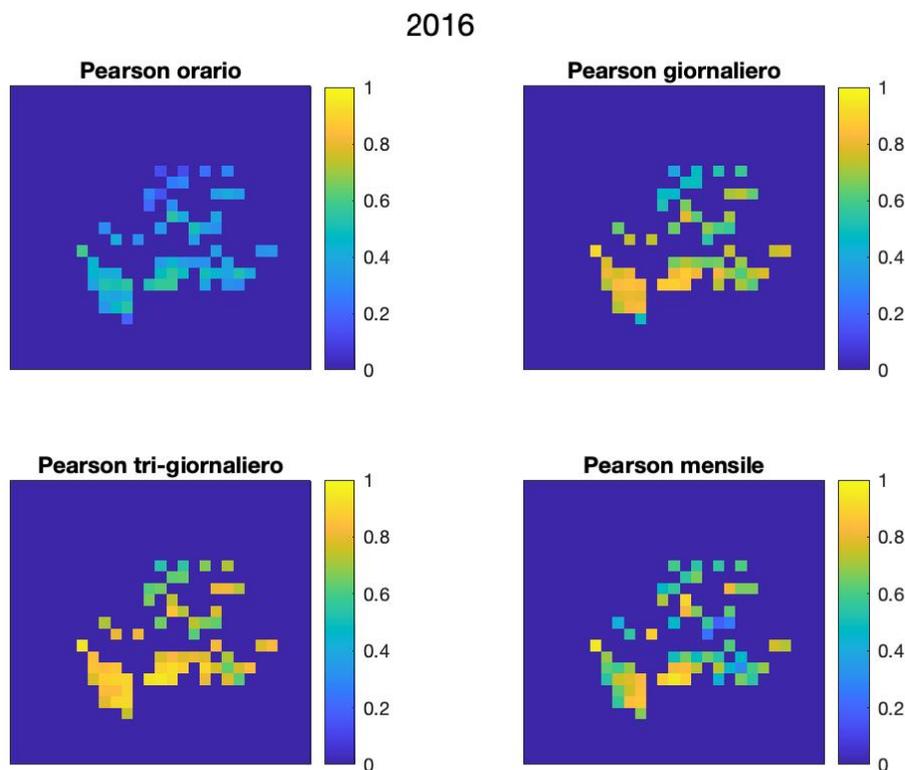


Figura 28 Distribuzione spaziale dell'indice di Pearson per l'anno 2016: allargando il periodo di aggregazione l'indice presenta valori maggiori, prossimi all'unità, evidenziando quindi una quasi perfetta corrispondenza tra variazioni della stima del sistema GPM e della misura pluviometrica. Si noti come le zone in cui l'indice risulta essere inferiore sono ancora quelle poste a nord-ovest, in un'area pedemontana, e quelle nella zona collinare a sud-est. Nell'area centrale del Piemonte, in pianura, l'indice presenta valori elevati.

Infine, consultando le altre rappresentazioni contenute in appendice (Figura 49, Figura 50, Figura 51, Figura 52), è possibile notare che le prestazioni differiscono notevolmente negli anni considerati. Durante il 2013 e il 2014 alle aggregazioni tri-giornaliere e mensili non corrispondono risultati particolarmente soddisfacenti, mentre nel 2016 e nel successivo 2017 invece contengono un elevato numero di caselle con indice elevato. La scala cromatica utilizzata rende di facile

individuazione questo fenomeno, in cui le mappe con coefficiente di Pearson basso tendono al blu, mentre prestazioni migliori si associano al giallo acceso passando brevemente per una tonalità verde acqua. Considerando periodi di aggregazione maggiori, il giallo-verde si impone come colore principale.

Si è scelto di rappresentare gli indici statistici tramite rappresentazioni grafiche di questo genere per individuare facilmente cluster di celle a prestazioni simili. Per completezza, occorre sottolineare che non tutte le mappe contengono lo stesso numero di celle, poiché come spiegato in precedenza si è scelto di escludere dai risultati tutte quelle celle che presentano valori mancanti in nella serie di dati orari.

Nonostante l'impiego di indici statistici e categorici possa semplificare l'interpretazione, l'osservazione diretta dei dati può fornire un'idea del comportamento dei due sistemi, in particolare il confronto tramite uno scatter plot e il calcolo del coefficiente di determinazione, o R^2 , permettono di valutare la variabilità dei dati e la correttezza della stima. In questo studio il valore di R^2 è stato calcolato tramite il quadrato del coefficiente di correlazione di Pearson, misurato tramite la funzione integrata in Matlab.

Si sono scelte dunque 5 caselle pseudocasuali distribuite omogeneamente sul territorio come in Figura 29, in azzurro su sfondo giallo, le cui coordinate vengono riportate a lato.

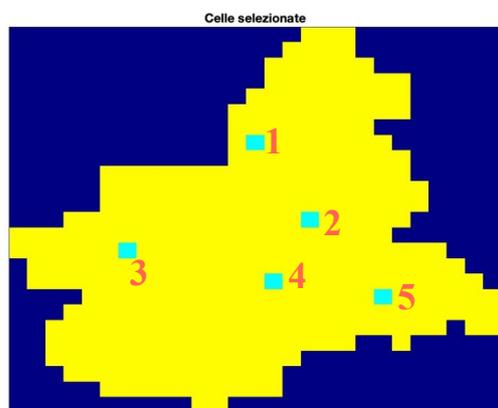


Tabella 6 Coordinate celle esempio, riferite al centro

N°	Lat.	Lon.
1	45,75°	7,95°
2	45,25°	8,25°
3	45,05°	7,25°
4	44,85°	8,05°
5	44,75°	8,65°

Figura 29 Celle di esempio scelte sulla base della posizione geografica, in modo da considerare in dettaglio la serie storica della precipitazione in aree strategiche.

Non verrà riportato l'intero insieme di serie storiche e gli scatter in ogni aggregazione, in quanto sono state ottenute oltre 3200 rappresentazioni.

Risulta interessante però indagare lo sviluppo della serie temporale di una cella che si ritiene possa essere in una zona con buona corrispondenza tra misura e stima, e una in cui la conformazione topografica è tale da causare possibili problemi di stima, come rispettivamente la cella indicata come numero 4 e posta in una zona di pianura (Figura 30, Figura 31, Figura 32, Figura 33) e la cella numero

1 sita in una zona collinare negli immediati pressi dei rilievi montuosi alpini (Figura 34, Figura 35, Figura 36, Figura 37), mentre le rappresentazioni relative alle altre celle saranno riportate in appendice. Guardando ancora al 2016, notiamo in entrambe le celle un miglioramento dell' R-quadro all'aumentare dell'aggregazione.

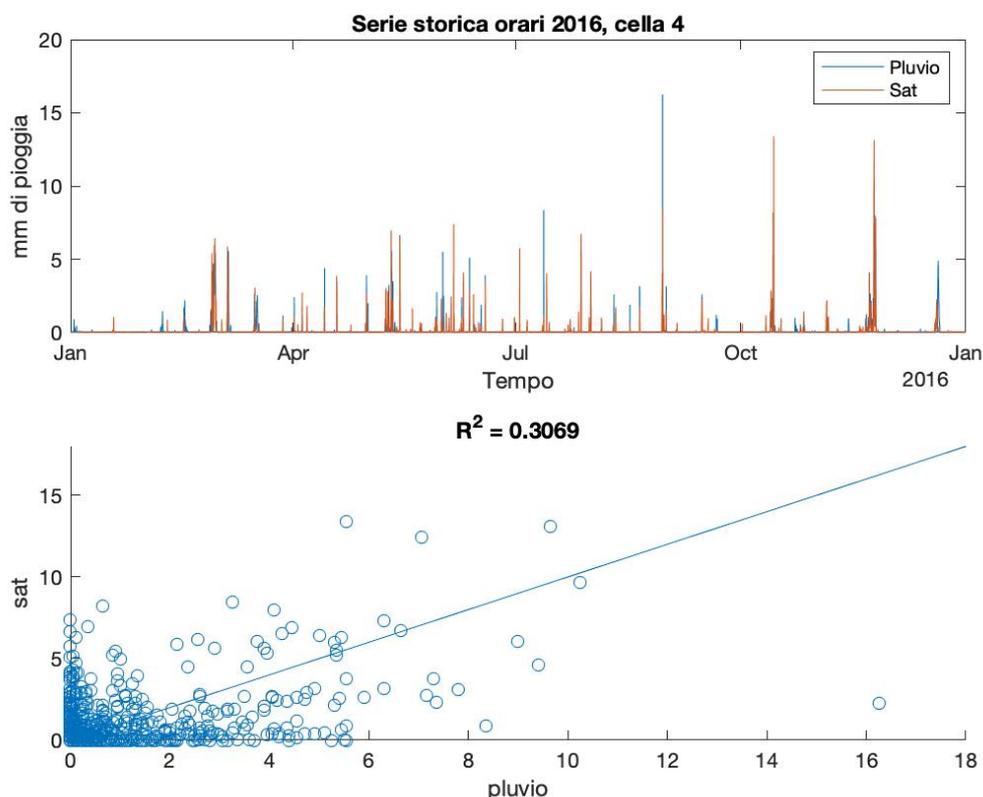


Figura 30 Serie storica e scatter, cella 4, valori orari. Sono presenti alcuni picchi in cui le due sorgenti sono discordi. Nello scatter plot la nuvola di punti è distribuita con valori anche distanti dalla linea di riferimento

Entrando maggiormente nello specifico, nella serie dei valori orari in Figura 30 possiamo osservare numerosi picchi in entrambe le fonti di dati, individuando in alcuni casi decise sovrastime da parte del sistema satellitare. Per semplicità interpretativa occorre riferirsi al sottostante scatter plot: notiamo come non siano molti i valori in perfetta corrispondenza allineati lungo la linea di riferimento, per la quale le due misure sarebbero uguali, inoltre è possibile notare numerosi punti in prossimità degli assi. Questi rappresentano uno zero per una delle due serie: i valori prossimi all'asse x una mancata misura satellitare e lungo l'asse y indicano un falso allarme. Nonostante ciò il risultato non è da considerare negativamente, non vi sono infatti eventi in cui la differenza di misura superi i 10mm di errore.

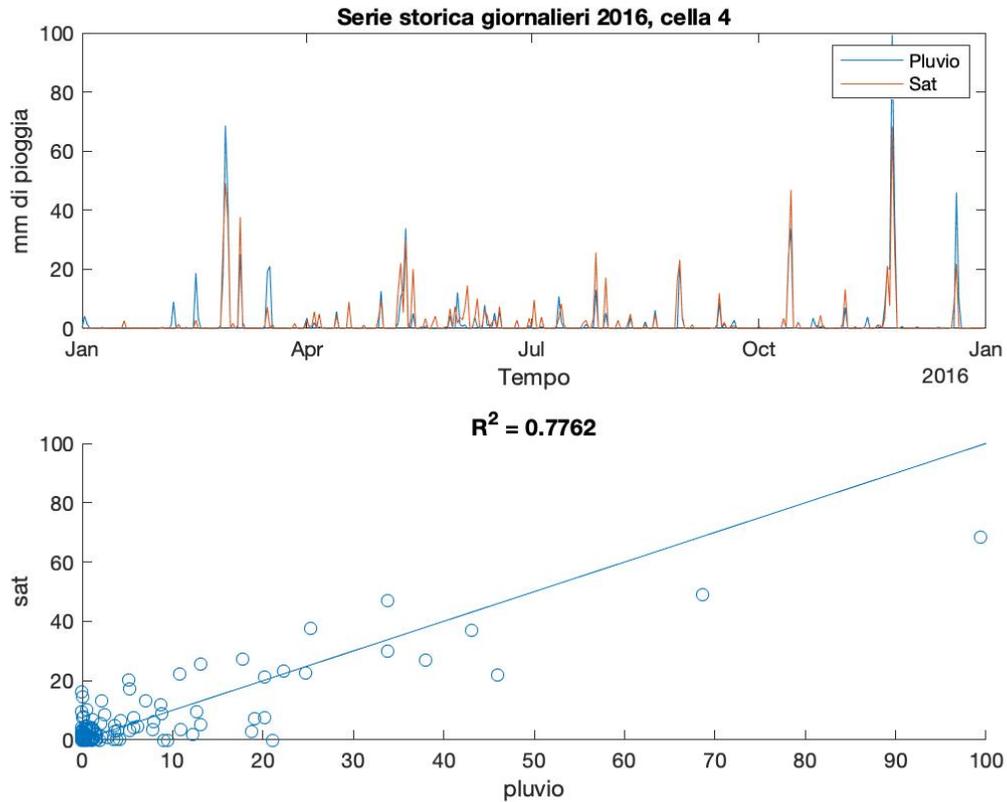


Figura 31 Serie storica e scatter, cella 4, valori giornalieri. I picchi in cui la stima non rispetta la precipitazione osservata sono in numero inferiore: si può infatti osservare una discreta accuratezza della stima, confermata dalla maggior vicinanza della nuvola di punti alla linea di riferimento nello scatter plot.

In Figura 31 è stata riportata la serie storica sotto forma di aggregazione giornaliera, in cui sono maggiormente evidenti gli eventi piovosi. Possiamo notare l'assenza di stime di precipitazione slegate dalla presenza di un evento osservato dalla rete pluviometrica, tuttavia la quantità stimata risulta scarsamente attigua alla misura pluviometrica in numerosi casi. In taluni, inoltre, IMERG produce una stima in eccesso, mentre per altri in difetto senza un'evidente correlazione con la quantità di pioggia o la stagionalità. Nello scatter plot i valori sono distribuiti su una fascia più ampia di valori, e la nuvola di distribuzione si avvicina maggiormente alla linea di riferimento: il valore del coefficiente di determinazione ha subito infatti un netto aumento rispetto a quello dello scatter plot delle misure orarie, passando da 0,3069 a 0,7762.

Si passa successivamente all'analisi dell'aggregazione successiva composta su 3 giorni di misure: similmente ai totali giornalieri, le stime di pioggia da parte del satellite si dimostrano capaci di rilevare correttamente nel tempo quando un fenomeno precipitativo è in atto, ma allargando la finestra temporale di acquisizione i millimetri di differenza con la misura effettuata dalla rete pluviometrica aumentano. Nello scatter plot di questa aggregazione tri-giornaliera

la stima satellitare e i dati pluviometrici presentano una maggior vicinanza: il valore di R^2 infatti aumenta maggiormente fino a 0,8563.

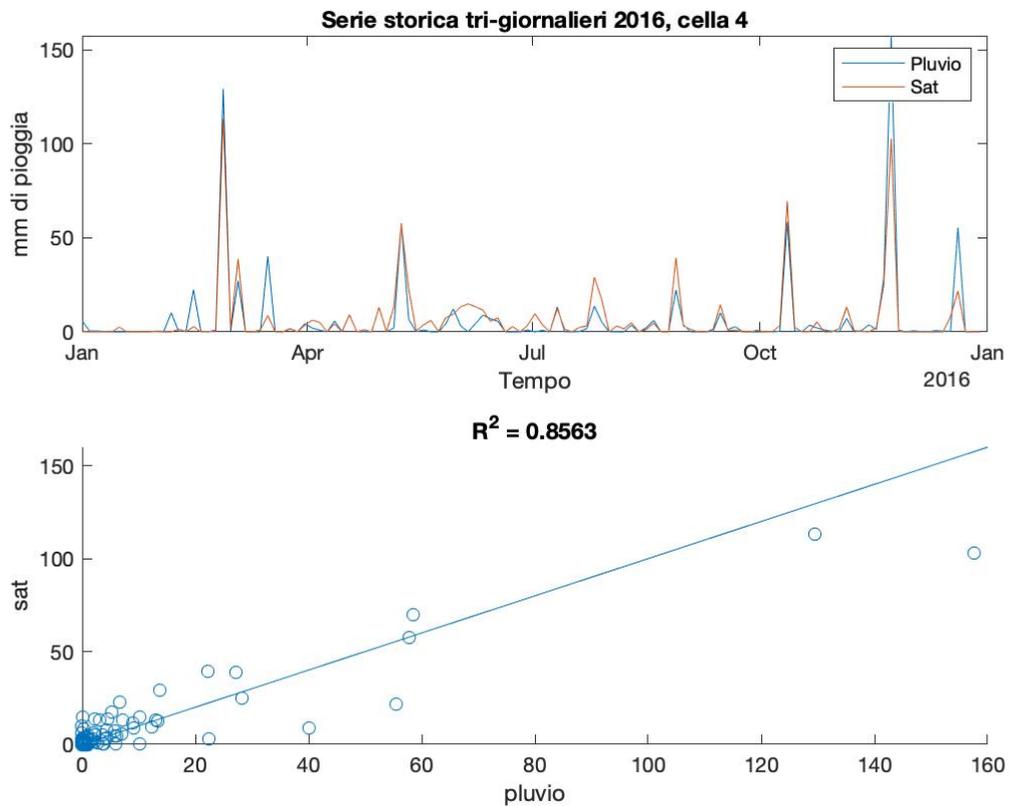


Figura 32 Serie storica e scatter, cella 4, valori per l'aggregazione tri-giornaliera. Vi è una discreta corrispondenza tra rete pluviometrica e stima da satellite: a conferma di ciò il coefficiente R^2 è elevato.

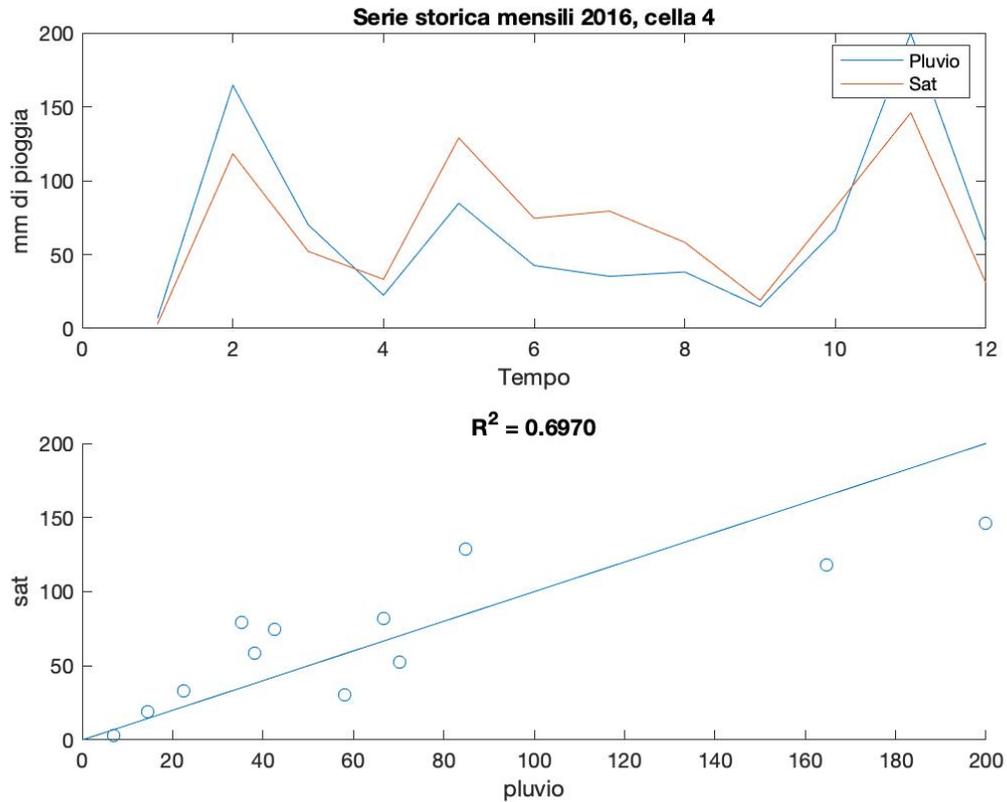


Figura 33 Serie storica e scatter, cella 4, valori mensili. L'andamento mensile osservato dalla rete a terra è rispettato dalla stima effettuata da GPM. Tuttavia, il coefficiente R^2 subisce un decremento a causa del minor numero di valori considerati e alle maggiori differenze in millimetri di pioggia dovute all'aggregazione stessa.

Consideriamo infine l'aggregazione mensile: risultano maggiormente evidenti le differenze quantitative tra le due sorgenti a causa dell'aggregazione su un periodo maggiore. È possibile notare un andamento concorde nella serie storica, ma a causa del limitato numero di valori considerati, pari al numero dei mesi in un anno, all'interno dello scatter plot la correlazione tra le due misure diminuisce e il coefficiente R^2 scivola a 0,6970.

Le ipotesi riguardo all'accuratezza dell'algoritmo IMERG nella produzione di una stima pluviometrica in un'area a bassa quota si possono quindi accettare come corrette: i valori di precipitazione presentano una discreta accuratezza confrontati con i dati raccolti dalle centraline a terra. Gli eventi principali vengono infatti riconosciuti come tali, così come la maggior parte di quelli di bassa-media entità, nonostante vi sia talvolta una scarsa precisione nella quantità di precipitazione stimata. Non è altresì emersa alcuna variazione stagionale dell'accuratezza.

Contrariamente a questa zona, nella cella indicata come numero 1, la situazione è opposta: il coefficiente di correlazione non supera in nessun caso il valore 0.35. Cominciando dall'osservazione della serie storica dei dati orari, questa appare all'osservatore come una frastagliata serie disconnessa di misure scollegate tra loro. A conferma di ciò, il relativo scatter plot evidenzia l'alto numero di stime

considerabili come “misses”, ovvero in cui vi è stata una precipitazione rilevata dal pluviometro ma il sistema satellitare non è stato in grado di rilevarla, ottenendo di conseguenza un coefficiente di determinazione di 0,01.

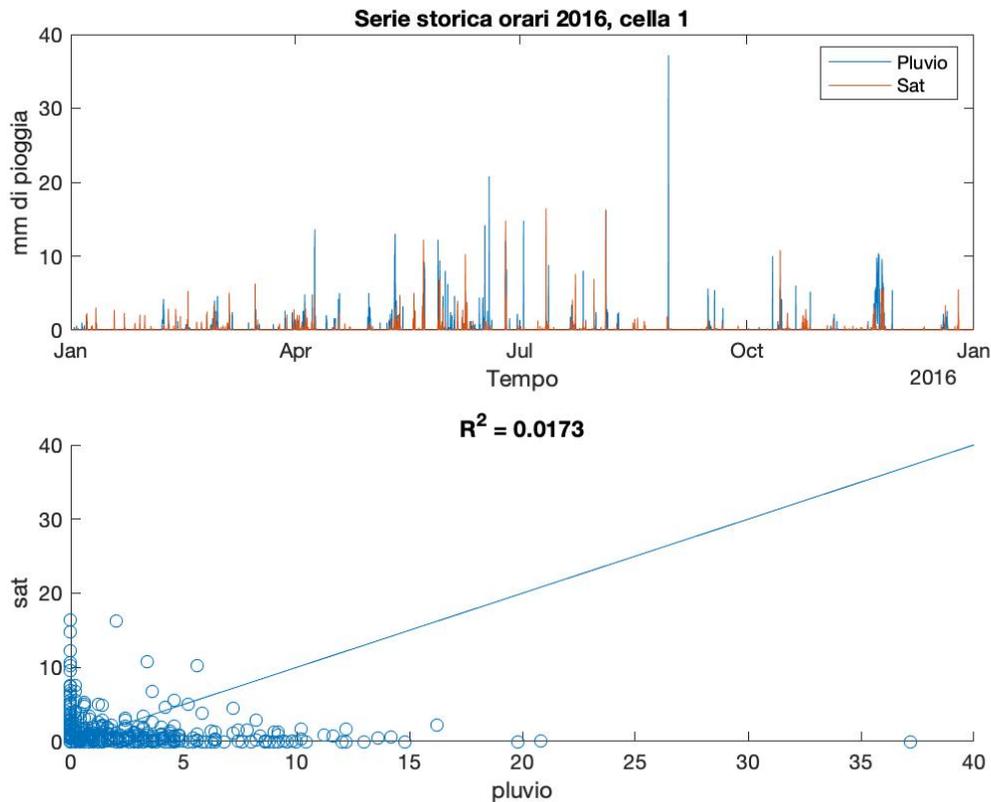


Figura 34 Serie storica e scatter, cella 1, valori orari. Sono presenti numerosi picchi in cui le due sorgenti sono discordi. Nello scatter plot la nuvola di punti è distribuita con valori anche distanti dalla linea di riferimento, con un R^2 molto basso.

Osservando la serie storica dell’aggregazione giornaliera risultano con maggior chiarezza le differenze tra rete pluviometrica e IMERG con differenze notevoli: ad esempio nel mese di dicembre un evento di precipitazione superiore a 120mm è stato stimato come inferiore ai 30mm. Il valore di R^2 è emblematico dello scarso miglioramento rispetto ai dati orari: 0,14.

Si consideri ora l’aggregazione di dati su 3 giorni: anche in questo caso appare lieve la correlazione tra stima e misura. È tuttavia possibile notare che le due misure presentano eventi piovosi in contemporanea, giustificando l’elevato valore degli indici categorici e il risultato parzialmente negativo degli indici statistici. R^2 possiede un valore di 0,28 e lo scatter plot si presenta come un insieme di punti scarsamente correlati.

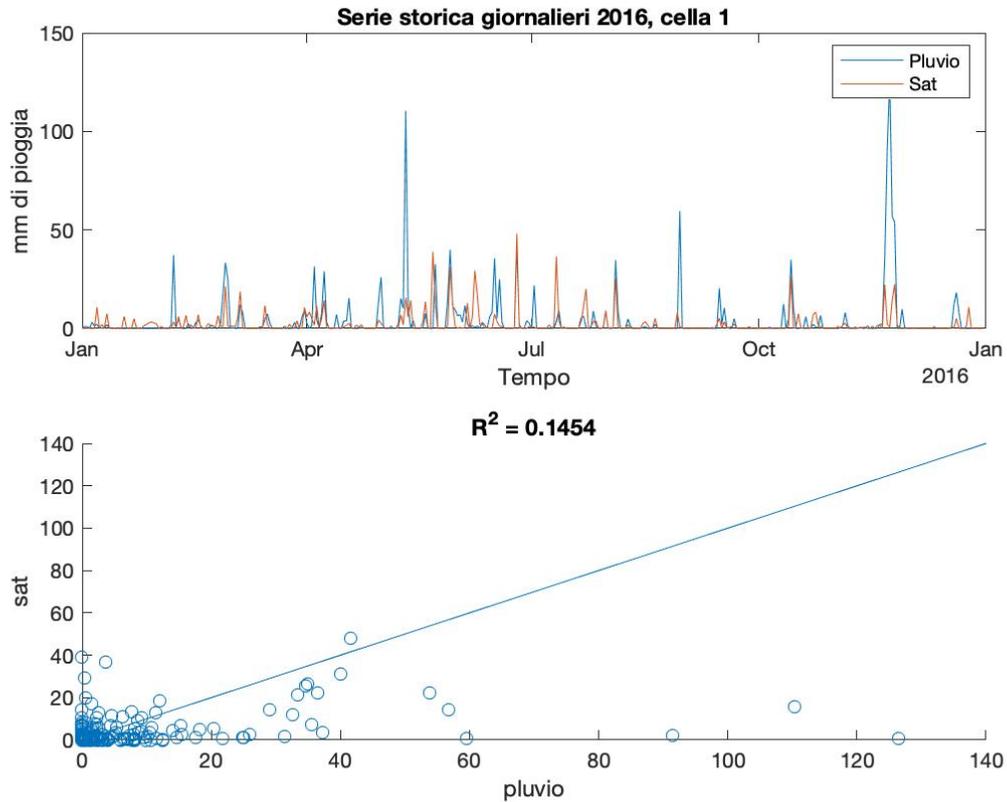


Figura 35 Serie storica e scatter, cella 1, valori giornalieri. I picchi in cui la stima non rispetta la precipitazione osservata sono in numero inferiore, tuttavia ve ne sono alcuni (ad esempio a maggio o novembre) in cui la differenza di stima è molto elevata. Il valore del coefficiente di determinazione rimane ridotto.

Infine, osserviamo l'aggregazione dei dati su base mensile: a causa della somma dei dati le differenze sono ancora maggiori, rendendo impossibile individuare un'evoluzione della serie storica in comune. Tuttavia, il valore del coefficiente di determinazione subisce un leggero aumento, diventando pari a 0,35, un valore rappresentativo di una scarsa accuratezza.

Risultati

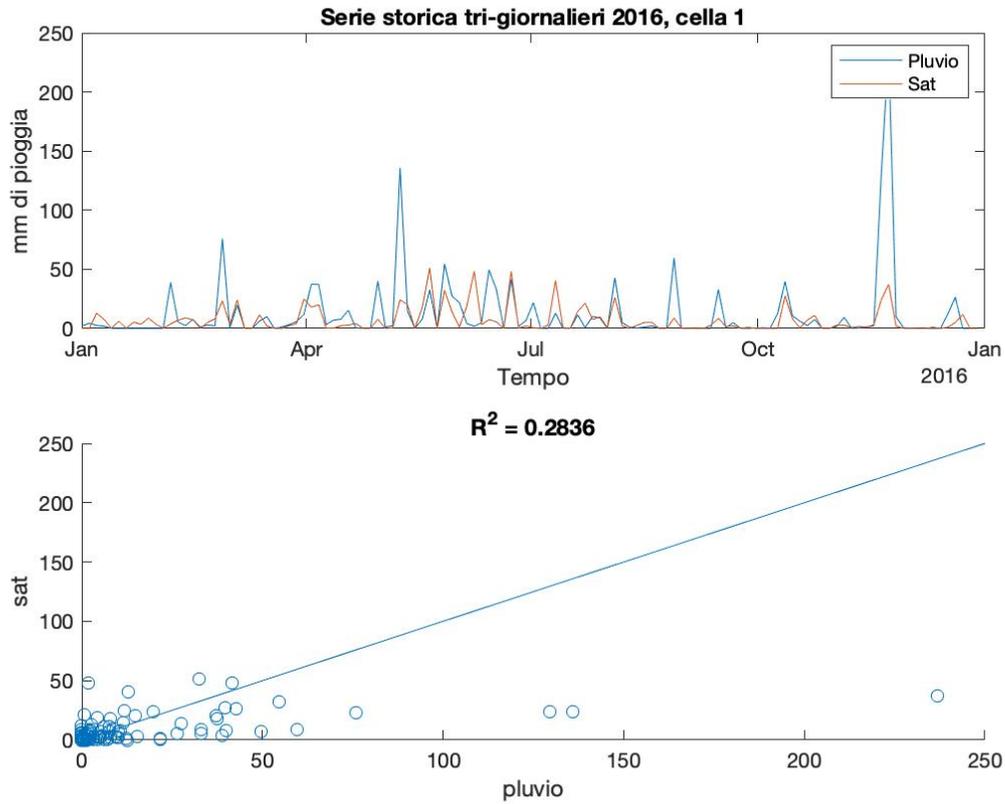


Figura 36 Serie storica e scatter, cella 1, valori per l'aggregazione tri-giornaliera. È possibile notare con chiarezza un periodo tra gennaio e febbraio dove in assenza di osservazione di precipitazione il sistema GPM ne ha stimato la presenza. Sono presenti numerose differenze tra stima e osservazione causando uno scatter plot in cui la nuvola di punti resta distante e molto ampia nei pressi della linea di riferimento.

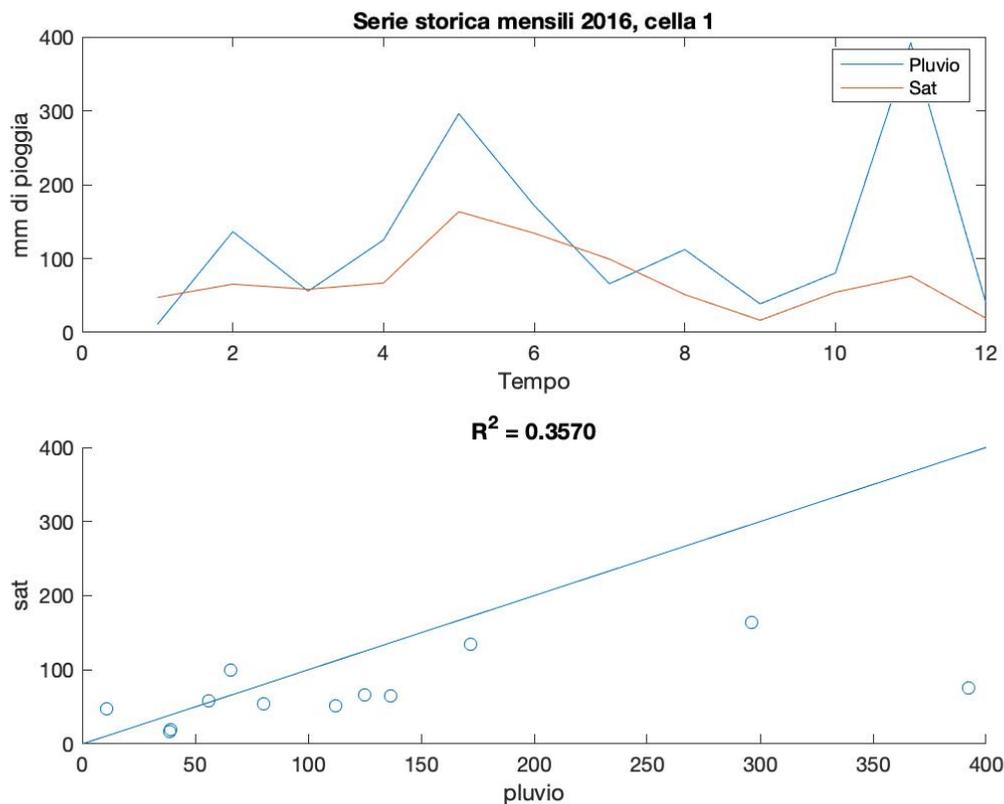


Figura 37 Serie storica e scatter, cella 1, valori mensili. L'andamento annuale della stima nei valori mensili appare molto differente dalla precipitazione osservata. Il valore di R^2 si mantiene molto ridotto indicando una scarsa accuratezza.

In Figura 38 e in Figura 39 è possibile vedere l'evoluzione negli anni delle prestazioni tramite il coefficiente di determinazione nelle celle indicate. Ricordando il cambio di strumentazione in orbita attorno alla Terra effettuato tra il 2014 e il 2015 inoltre, non è possibile individuare alcun miglioramento nella qualità di stima dell'algorithmo di IMERG.

Risultati

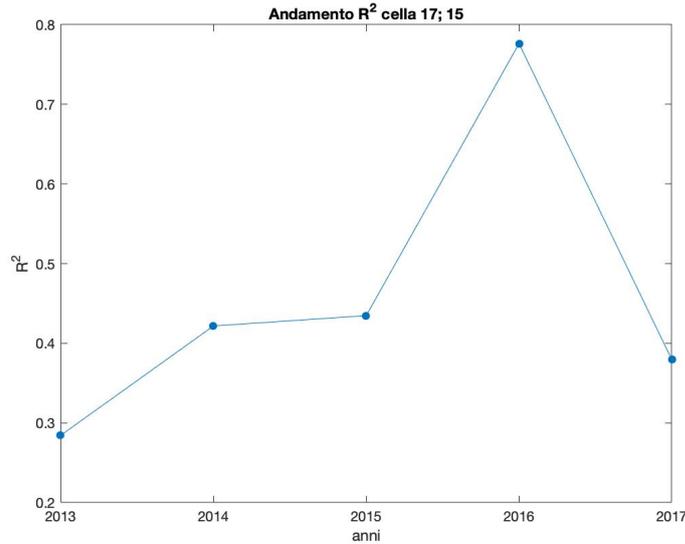


Figura 38 Variazione coefficiente di determinazione per l'aggregazione giornaliera, cella 4. Il 2016 è l'anno in cui la prestazione è migliore e la stima è maggiormente simile alla realtà osservata.

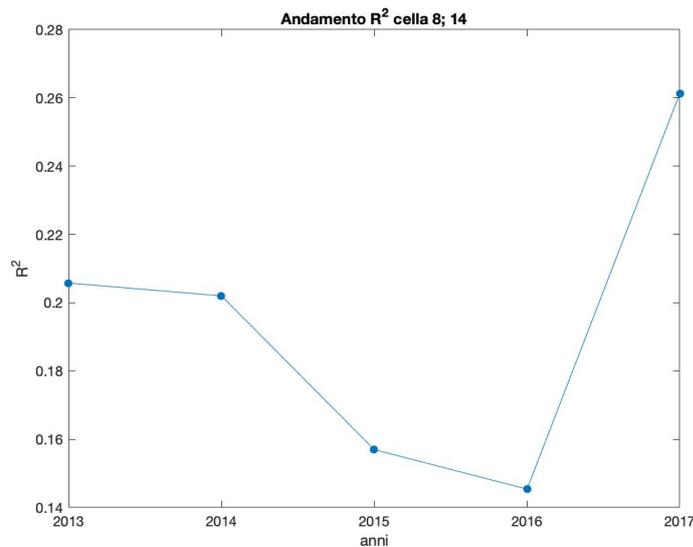


Figura 39 Variazione coefficiente di determinazione, cella 1. In tutti gli anni il valore di R^2 è tale da considerare negativamente la stima da parte di IMERG. Il 2016 è l'anno peggiore, ma il coefficiente è ridotto in tutti gli anni considerati.

Per poter effettuare un'ulteriore valutazione dell'accuratezza della stima possiamo guardare alla variazione spaziale del coefficiente di determinazione, sia nelle differenti aggregazioni temporali in Figura 40, che durante i restanti anni considerati all'interno di questo studio, in appendice. Si è scelto di riportare l'anno 2015 poiché permette di evidenziare meglio la variazione del coefficiente R^2 . Osservando la distribuzione nei dati orari notiamo un valore di R^2 generalmente poco elevato, inferiore a 0,4, per aumentare successivamente nell'aggregazione

giornaliera superando in alcuni casi 0,6. Nell'aggregazione di 3 giorni aumentano le celle in cui il coefficiente di determinazione ha valori superiori a 0,6, ma nella zona nei pressi dell'arco alpino in tutta la fascia settentrionale-ovest del Piemonte il coefficiente si mantiene su valori inferiori. Questa caratteristica distribuzione si mantiene nell'aggregazione mensile, in cui è altresì possibile osservare isolati valori in cui R^2 è maggiore di 0,8 avvicinandosi molto all'unità.

In appendice (Figura 53, Figura 54, Figura 55, Figura 56) si trovano le rappresentazioni degli altri anni per cui sono stati calcolati gli indici di valutazione: notiamo come il 2016 sia l'anno in cui la stima si avvicina maggiormente alla misura effettuata dalla rete pluviometrica, includendo ampie zone ad elevato coefficiente di determinazione. Al contrario, gli anni 2013 e 2014 presentano valori nettamente inferiori: possiamo attribuire questo comportamento alla strumentazione meno recente appartenente alla missione TRMM. In Tabella 7 si trova uno specchietto riepilogativo del valore di R^2 medio nei differenti anni ed aggregazioni.

Tabella 7 Indice di determinazione R^2 per differenti anni e tempi di aggregazione

	2013	2014	2015	2016	2017	Media
Orari	0,09	0,12	0,11	0,15	0,08	0,11
Giornalieri	0,29	0,37	0,32	0,50	0,32	0,36
Tri-giornalieri	0,38	0,41	0,39	0,62	0,37	0,43
Mensili	0,19	0,39	0,33	0,44	0,39	0,35

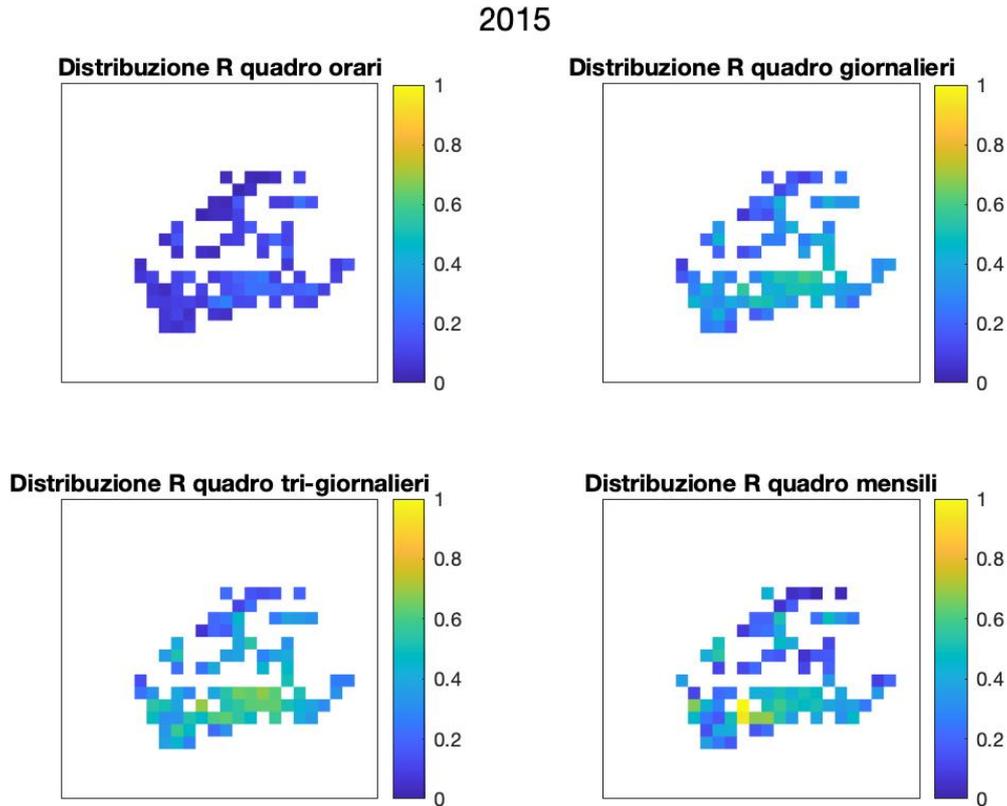


Figura 40 Distribuzione spaziale del coefficiente di determinazione per anno 2015. All'aumentare dell'aggregazione notiamo un miglioramento dell'indice. Le celle a nord della regione presentano i valori peggiori, mentre nell'area pianeggiante a sud-ovest si trova un'area in cui R^2 è generalmente al di sopra di 0,6.

Considerando complessivamente la distribuzione spaziale degli indici statistici, RMSE e coefficiente di Pearson, del coefficiente di determinazione R^2 , e le serie storiche riportate per celle in posizioni ritenute rappresentative, è possibile affermare che la capacità della missione GPM di produrre una stima accurata di precipitazione è ottimale in aree pianeggianti o con una ridotta elevazione, mentre in zone poste nei pressi di rilievi montuosi o a vocazione collinare la stima è soggetta a forti cali di accuratezza.

5. Discussione dei risultati

A seguito della presentazione dei principali risultati ottenuti dal confronto dell'osservazione della precipitazione da parte di una rete pluviometrica e della stima della stessa grandezza effettuata da un sistema satellitare, è emersa una discreta correlazione tra i risultati, con un coefficiente di determinazione R^2 medio di 0,36 nell'aggregazione giornaliera dei dati e di 0,43 in quella composta su 3 giorni, nonostante vi siano zone in cui questo valore risulta essere nettamente maggiore. Tuttavia, dall'analisi emerge che non vi sia la capacità, da parte dei dati prodotti dall' algoritmo IMERG, di sostituire una misura operata in campo. In Tabella 8 è mostrata la percentuale di successo nel rilevare la pioggia correttamente, data dalla somma dei valori classificati come *correct negatives* e *hits*, ovvero valori nulli in assenza di precipitazione e valori positivi quando questa è stata osservata. Notiamo che il sistema è in grado di acquisire con successo la presenza precipitazione in almeno il 70% dei casi nell'aggregazione giornaliera. Tuttavia, la precisione quantitativa risente della complessità del fenomeno precipitativo, presentando valori di RMSE medio sull'aggregazione giornaliera pari a 6,6 mm.

Tabella 8 Percentuali di successo dell'algoritmo IMERG nella rilevazione della precipitazione, divise per aggregazione temporale e anno di riferimento

ANNO	Orari	Giornalieri	Tri_giornalieri	Mensili
2013	0,91	0,7093866	0,713484864	1
2014	0,9	0,7225998	0,741735744	1
2015	0,93	0,73693	0,683061184	0,8929
2016	0,92	0,7395322	0,695660467	0,9748
2017	0,94	0,7709929	0,694545455	0,9163

In altre parole, la capacità di rilevare la *presenza* di precipitazione è buona, mentre la capacità di rilevarne la *quantità* è scarsa. Troviamo conferma di ciò osservando ancora i risultati raggiunti: in particolare nell'aggregazione giornaliera, considerata maggiormente rappresentativa, si può notare che la probabilità di rilevamento (POD) è di poco inferiore o pari al 60%. Contemporaneamente però, il RMSE sui valori giornalieri di ogni cella è spesso superiore ai 3-5mm, un quantitativo di acqua che, per applicazioni di precisione alle latitudini piemontesi, può influenzare pesantemente la corretta gestione dell'irrigazione e dei trattamenti.

Dallo studio è emersa anche una relazione tra precisione della stima satellitare ed altitudine media della cella: nelle zone pedemontane ai limiti della rete pluviometrica, in particolar modo nella zona del biellese e nella zona collinare a

sud della città di Alba, non è stato infrequente individuare aree in cui tutti gli indici hanno presentato risultati negativi rispetto alle altre zone considerate.

Di seguito si elencheranno le criticità principali e possibili estensioni di questo studio.

- Il confronto si è limitato alle sole celle in cui vi sono stazioni meteorologiche dotate di pluviometro
 - Sarebbe possibile spazializzare i dati in un'area di interesse definita a priori tramite metodi di interpolazione quali IDW o kriging. Studi precedenti hanno avuto esiti interessanti applicando la spazializzazione kriging nella validazione di GPM. (Goudenhoofdt 2009) (Xu 2016) (Yan 2019) (Chiaravalloti 2018) (Brommundt 2007)
- Gli intervalli delle aggregazioni temporali sono stati scelti principalmente sulla base dell'utilità della misura di precipitazione per applicazioni di precisione, ovvero dati orari e giornalieri hanno un'elevata importanza, mentre le aggregazioni mensili sono necessarie per monitorare l'andamento annuale.
 - È possibile includere numerosi altri periodi di aggregazione, ad esempio basati su 14 giorni, oppure valutare la prestazione su base stagionale.
- Sono state escluse le celle in cui vi sono dati mancanti nella serie storica
 - Effettuare la stessa analisi con serie storiche incomplete porterebbe ad ottenere risultati soggetti a margini di errore maggiori, ma comunque validi nel caso di isolate mancanze di valori orari.

6. Conclusioni

Al termine di questo studio è possibile affermare che la stima di precipitazione proposta da GPM ha numerosi aspetti positivi: oltre alla già citata valenza a scala mondiale, rispecchia l'andamento mensile medio, raggiunge indici prestazionali discreti, l'intervallo di tempo che intercorre tra acquisizione e pubblicazione del dato è ridotto, ed infine lo studio di validazione viene effettuato a livello mondiale e confermato da numerosi studi indipendenti.

Tuttavia, non si ritiene che tale stima possa essere uno strumento sostitutivo alle misure effettuate tramite pluviometri per tre motivi principali:

1. Per applicazioni di precisione l'accuratezza non è sufficiente: in media la stima ha uno scarto di 6,6 mm rispetto alla misura pluviometrica.
2. IMERG propone il solo valore di precipitazione elaborato. Altri parametri, definiti come ausiliari, che includono ad esempio temperatura, umidità, pressione, radiazione solare, ecc, sono utilizzati per produrre la stima di pioggia, ma non sono disponibili all'utente. Tuttavia, nel caso si renda necessario effettuare un monitoraggio ambientale, il solo parametro di pioggia non è sufficiente per condurre un'analisi completa.
3. Il controllo sull'algoritmo è limitato: non sono infatti note le operazioni effettuate per ottenere la stima di precipitazione, per cui l'utente ha scarsa tracciabilità del processing del dato.

Le tre motivazioni sopra elencate, secondo le osservazioni attuate durante la stesura di questo studio, sono circoscritte alla possibilità di utilizzare l'output di IMERG in sostituzione ad una rete pluviometrica. Sul territorio considerato, in ambito agricolo, purtroppo il sistema satellitare non rappresenta un'alternativa equivalente.

Tuttavia, vi è la possibilità di sfruttare con profitto questo metodo di stima in numerose altre applicazioni, ove non sia disponibile una rete pluviometrica locale sufficientemente capillare. Esempi di applicazioni alternative possono essere i seguenti: impiego in zone remote in cui il mantenimento di uno strumento a terra risulterebbe eccessivamente oneroso, oppure l'utilizzo come elemento integrativo al fine di spazializzare il dato registrato puntualmente. Il sistema GPM può anche essere un ottimo "pioniere", o un sostituto temporaneo di un pluviometro, di cui l'utente deve conoscere limiti e capacità. Sono, infatti, già attive applicazioni emergenziali, in cui IMERG è utilizzato per monitorare in near-real-time l'evoluzione di eventi estremi per produrre previsioni in continuo aggiornamento. Un possibile proseguimento di questo studio è da ricercare nell'implementazione di una calibrazione dell'algoritmo di stima, per renderlo maggiormente performante in un contesto geografico complesso come l'area presa in esame.

Bibliografia

- Annan, Kofi. 2018. *kofiannanfoundation.org*. den 25 Giugno.
<https://www.kofiannanfoundation.org/articles/sustainable-water/>.
- Asong, et al. 2017. "Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over Southern Canada against Ground Precipitation Observations: A Preliminary Assessment." *American Meteorological Society*.
- Brommundt , et al. 2007. "Spatial correlation of radar and gauge precipitation data in high temporal resolution." *Elsevier B.V.*
- Caracciolo, et al. 2018. "Performances of GPM satellite precipitation over the two major T Mediterranean islands." *Elsevier B.V.*
- Chen, et al. 2016. "Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China." *Remote Sensing*.
- Chiaravalloti, et al. 2018. "Assessment of GPM and SM2RAIN-ASCAT rainfall products over complex terrain in southern Italy." *Elsevier B.V.*
- Colli, M., L.G. Lanza, och P. La Barbera. 2013. "Performance of a weighing rain gauge under laboratory simulated time-varying reference rainfall rates." *Atmospheric Research*.
- Deutscher wetterdienst. u.d. Använd 2020. www.dwd.de/.
- Deutscher Wetterdienst. u.d.
https://opendata.dwd.de/climate_environment/GPCC/PDF/GPCC_intro_products_v2018.pdf.
- ECMWF. 2020. Använd 2020.
<https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation>.
- Encyclopædia Britannica. 2019. <https://www.britannica.com/science/weather-forecasting/Application-of-radar#ref943209>. Använd 2020.
- ESRI. 2019. (<https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm>).
- Gjertsen, et al. 2004. "Gauge adjustment of radar-based precipitation estimates in Europe."
- Goudenhoofd, et al. 2009. "Evaluation of radar-gauge merging methods for quantitative precipitation estimates."
- Habib, E., W.F. Krajewski, och A. Kruger. 2001. "Sampling errors of tipping-bucket rain gauge measurements." *Journal of Hydrologic Engineering*.
- Jah, Alok. 2016. *The Water Book*. Londra: Headline Publishing Group.
- Luca Fornaciari Fotografia. u.d. Använd 2020.
<https://www.lucafornaciarifotografia.com/>.
- Mahadi Hasan, et al. 2015. "Merging radar and in situ rainfall measurements: An assessment of different combination algorithms."
- Mazzoglio, et al. 2019. "Improving an Extreme Rainfall Detection System with GPM IMERG data." *Remote Sensing*.
- Meadows, Donella H. , D.L. Meadows, J. Randers, och W.W. Behrens III. 1972. *The Limits to Growth*. Potomac Associates.
- NASA. 2015. Använd 2020.
https://www.nasa.gov/mission_pages/GPM/spacecraft/index.html.
- . 2015. Använd 2020. <https://pmm.nasa.gov/TRMM/>.
- . 2015. <https://pmm.nasa.gov/GPM/flight-project/DPR>.
- . 2015. <https://gpm.nasa.gov/GPM/flight-project/core-observatory>.
- . 2020. <https://gpm.nasa.gov/data-access/downloads/gpm>.
- . 2020. <ftp://jsimpson.pps.eosdis.nasa.gov/NRTPUB/imerg/early/>.

- Nasa Algorithm Theoretical Basis Document V06. 2019.
- Nashwan, et al. 2019. "Assessment of Satellite-Based Precipitation Measurement Products over the Hot Desert Climate of Egypt." *Remote Sensing*.
- NOAA - National Oceanic and Atmosphere Administration. u.d. www.weather.gov. Använd 2020. https://www.weather.gov/bmx/radar_dualpol.
- NOAA, AP/Wide World Photos . u.d. *National Oceanic and Atmospheric Administration*;
- radartutorial.eu. u.d. <https://www.radartutorial.eu/01.basics/Radar%20Coverage.en.html>. Använd 2020.
- Regione Piemonte. 2019. www.regione.piemonte.it. Använd 2020. <https://www.regione.piemonte.it/web/temi/agricoltura/agroambiente-meteosuoli/rete-agrometeorologica-piemonte-ram>.
- Remote Sensing Systems. u.d. Använd 2020. <http://www.remss.com/measurements/rain-rate/>.
- Sevruk, B., och L. Zahlavova. 1994. "Classification system of precipitation gauge site exposure: Evaluation and application." *International Journal of Climatology*.
- Sevruk, B., och V. Nespor. 1999. "Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation." *Journal of Atmospheric and Oceanic Technology*.
- United Nations. 2015. Använd 2020. <https://sdgs.un.org/goals>.
- Willmott, C.J., och D.R. Legates. 1990. "Mean seasonal and spatial variability in gauge-corrected, global precipitation." *Int. J. Climatology*.
- World Meteorological Organisation. 2017. Använd 2020. https://library.wmo.int/doc_num.php?explnum_id=3152.
- Xu, et al. 2016. "Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network."
- Yan, et al. 2019. "Short time precipitation estimation using weather radar and surface T observations: With rainfall displacement information integrated in a stochastic manner."

Appendice

Si riportano in appendice i diagrammi non esplicitamente trattati nel testo principale, ritenuti tuttavia di importanza ai fini della completezza dello studio.

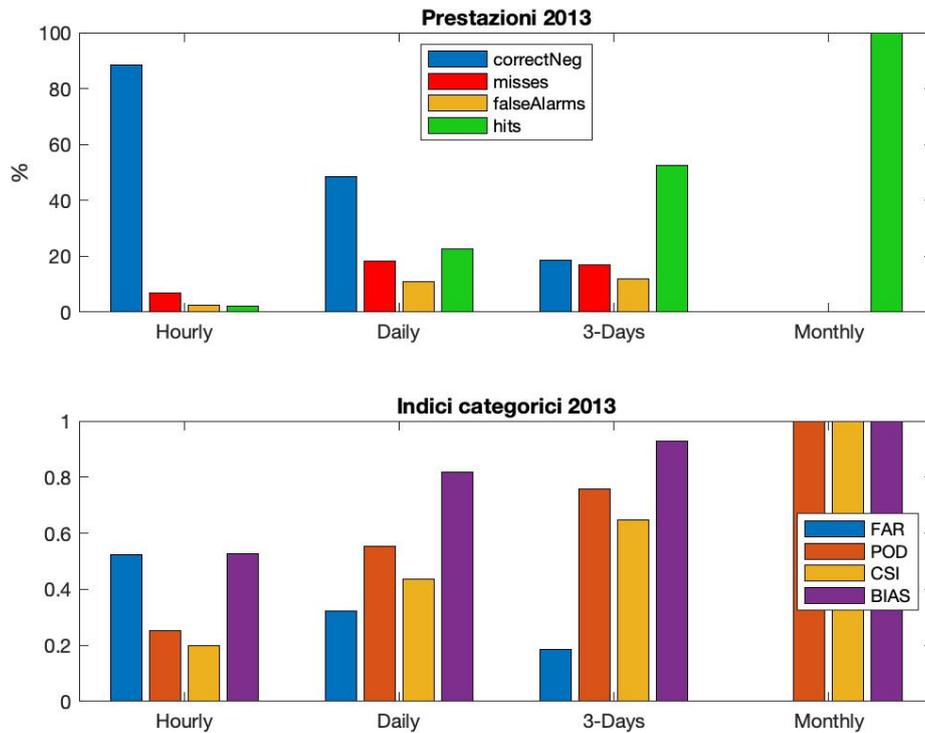


Figura 41 Risultati della tabella di contingenza per l'anno 2013: all'aumentare della finestra temporale sull'asse delle ascisse diminuisce la probabilità di incorrere in periodi senza precipitazione, per cui i correct negatives subiscono una riduzione; contemporaneamente gli hits aumentano, poiché è sufficiente rilevare una singola precipitazione all'interno del periodo di aggregazione per classificare come corretta la stima. Al di sotto invece si trova l'evoluzione degli indici categorici al variare del periodo di aggregazione: il valore ottimale è pari a 0 per il FAR e pari ad 1 per gli altri indici, evidenziando un comportamento positivo della stima.

Appendice

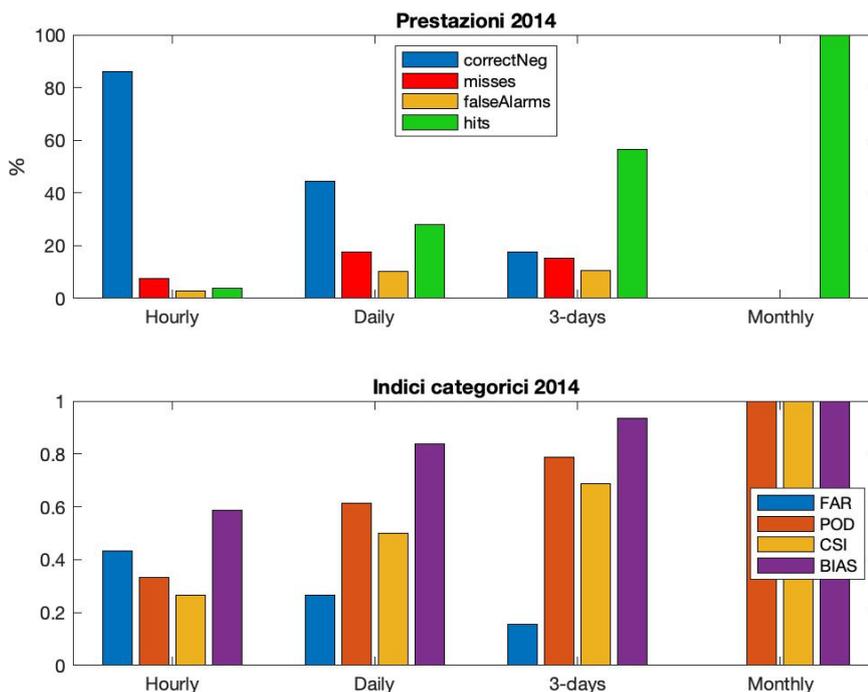


Figura 42 Risultati della tabella di contingenza 2014: l'evoluzione degli indici categorici e delle categorie prestazionali al variare del periodo di aggregazione temporale rispecchia il comportamento dei medesimi indici nell'anno precedente.

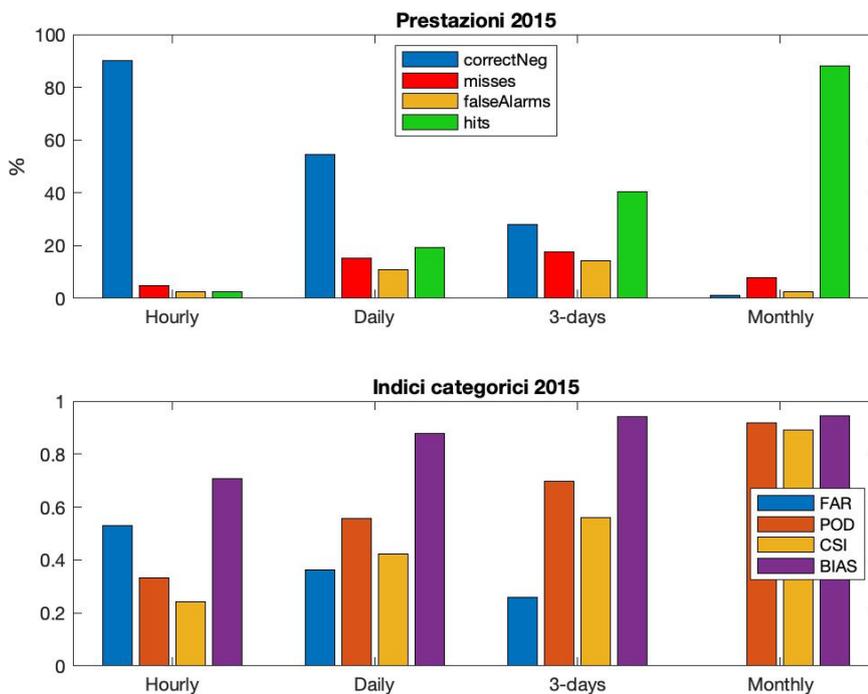


Figura 43 Risultati della tabella di contingenza 2015: l'evoluzione degli indici categorici e delle categorie prestazionali al variare del periodo di aggregazione temporale rispecchia il comportamento dei medesimi indici nell'anno precedente. Tuttavia, è possibile notare un minor successo negli indici categorici, evidente nei valori percentuali inferiori nell'aggregazione mensile.

Appendice

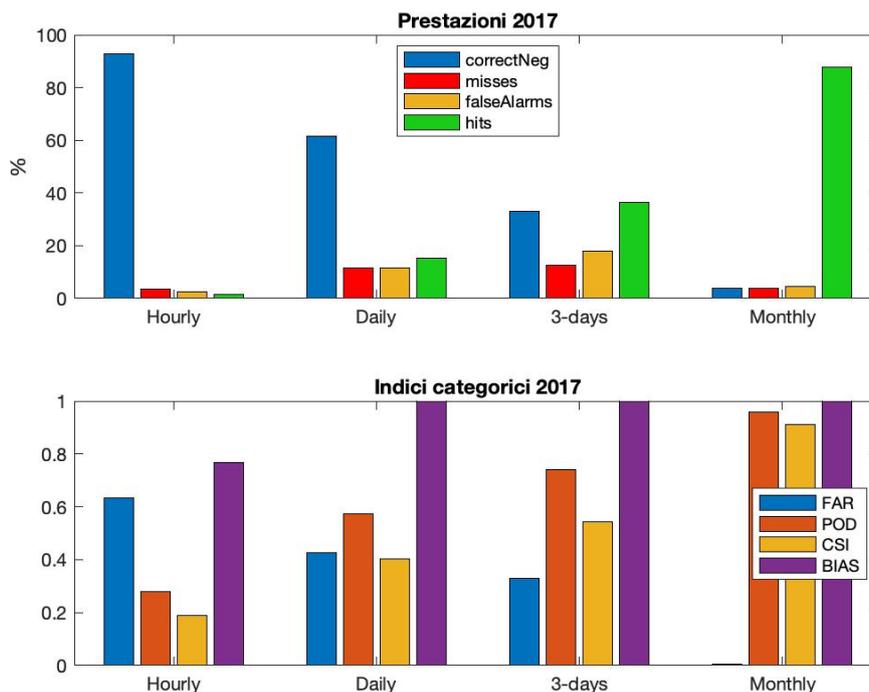


Figura 44 Risultati della tabella di contingenza 2017: l'evoluzione degli indici categorici e delle categorie prestazionali al variare del periodo di aggregazione, a differenza degli anni precedenti, presenta una maggior percentuale di false alarms. Tuttavia, in generale il comportamento è tale da poter essere accostato a quello degli altri anni considerati nell'analisi.

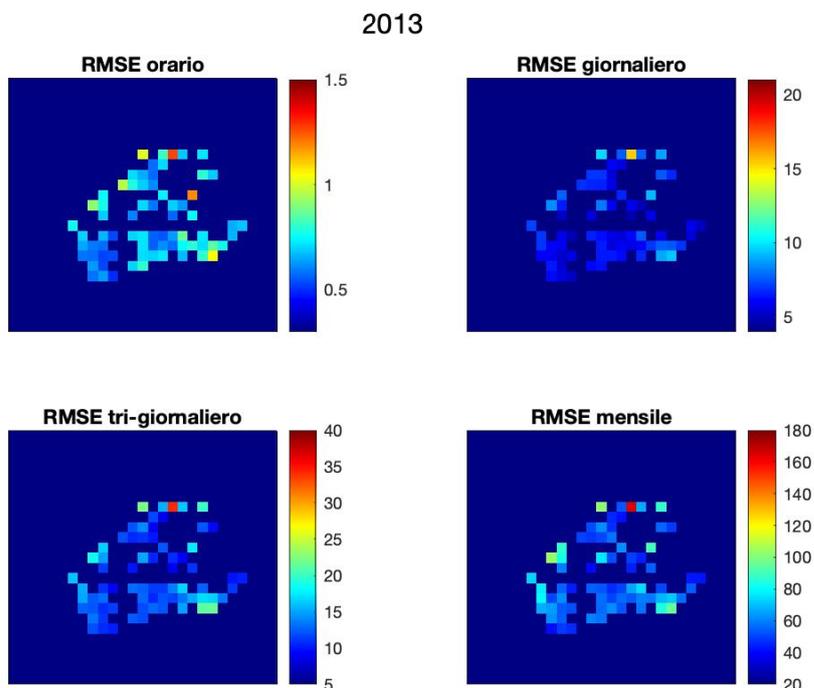


Figura 45 Distribuzione spaziale dei valori di RMSE 2013: i valori aumentano dal blu al rosso. I valori maggiori di scarto si trovano nelle zone a nord ed a ovest, in prossimità dell'arco alpino, dove le stazioni incluse nella cella hanno un'altitudine maggiore. Vi è inoltre una piccola area a sud in cui la topografia è collinare e lo scarto risulta maggiore. Al contrario, al centro della distribuzione, dove prevale la pianura, i valori di scarto sono nettamente inferiori.

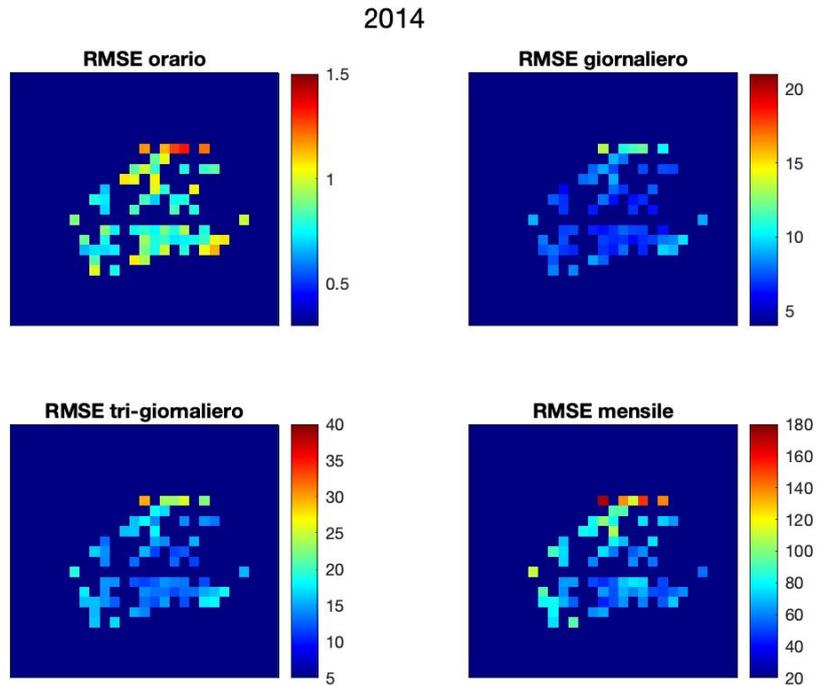


Figura 46 Distribuzione spaziale dei valori di RMSE 2014: si può notare la presenza di valori maggiori nelle aree limitrofe a nord, ovest e sud-est, in cui la vicinanza all'arco alpino e la topografia di natura collinare produce errori di stima da parte di IMERG.

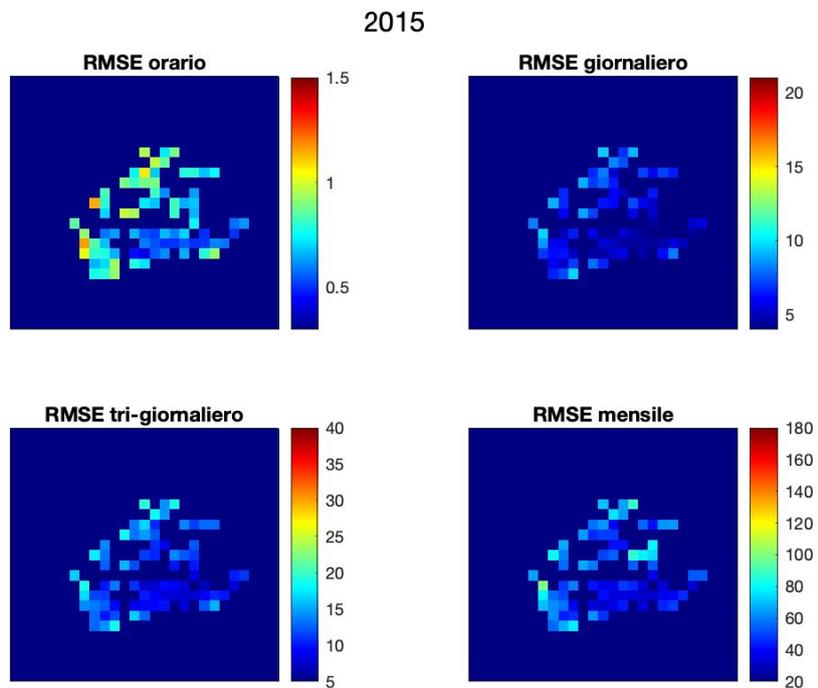


Figura 47 Distribuzione spaziale dei valori di RMSE 2015: rispetto all'anno precedente i valori di scarto si presentano minori, con ampie aree nelle quali RMSE si mantiene inferiore a 10mm nelle aggregazioni giornaliere e tri-giornaliere.

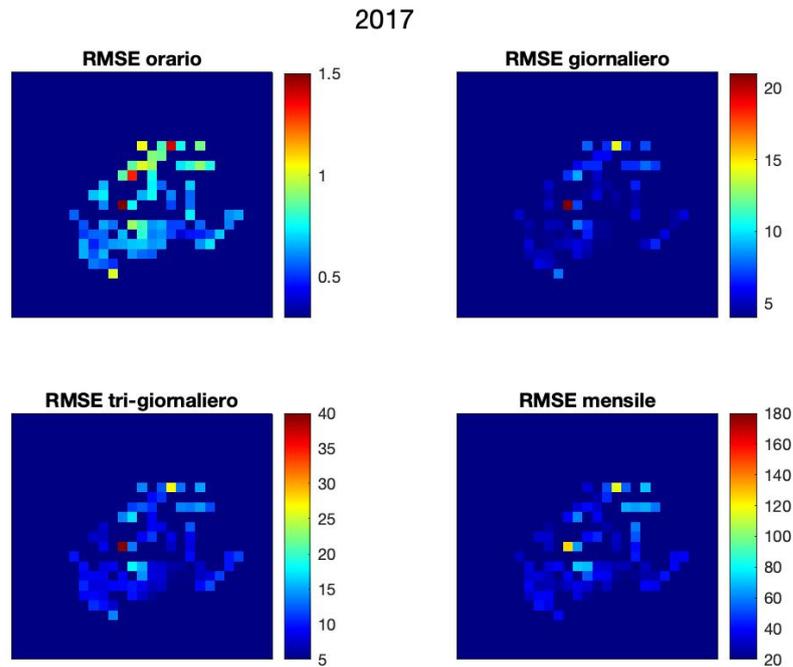


Figura 48 Distribuzione spaziale dei valori di RMSE per l'anno 2017: è possibile osservare valori di scarto inferiori agli anni precedenti, in particolare rispetto al periodo 2013-2015. Un possibile spiegazione può essere attribuita al definitivo utilizzo di satelliti della missione GPM ed all'abbandono di quelli di competenza di TRMM. Anche in questo caso, tuttavia, la stima di pioggia presenta uno scarto elevato nella zona settentrionale della distribuzione, mentre al centro si nota un solo valore anomalo.

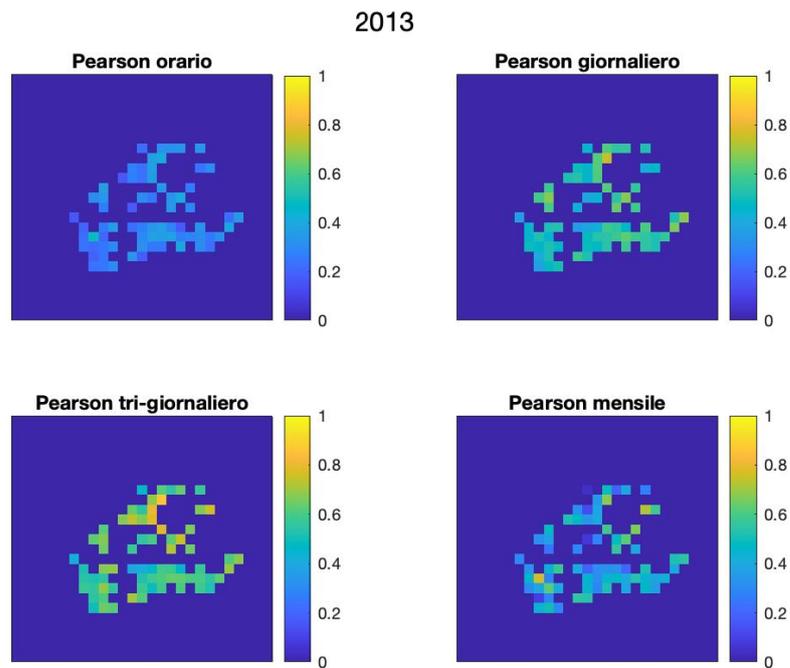


Figura 49 Distribuzione spaziale dell'indice di Pearson 2013: aumentando il periodo di aggregazione, l'indice presenta valori maggiori, prossimi all'unità in taluni casi, evidenziando quindi una quasi perfetta corrispondenza tra variazioni della stima del sistema GPM e della misura pluviometrica. Si noti come le zone in cui l'indice risulta essere inferiore sono ancora quelle poste a nord-ovest, in un'area pedemontana, e quelle nella zona collinare a sud-ovest. L'indice presenta valori maggiori nelle aggregazioni di 3 giorni rispetto a quelle mensili.

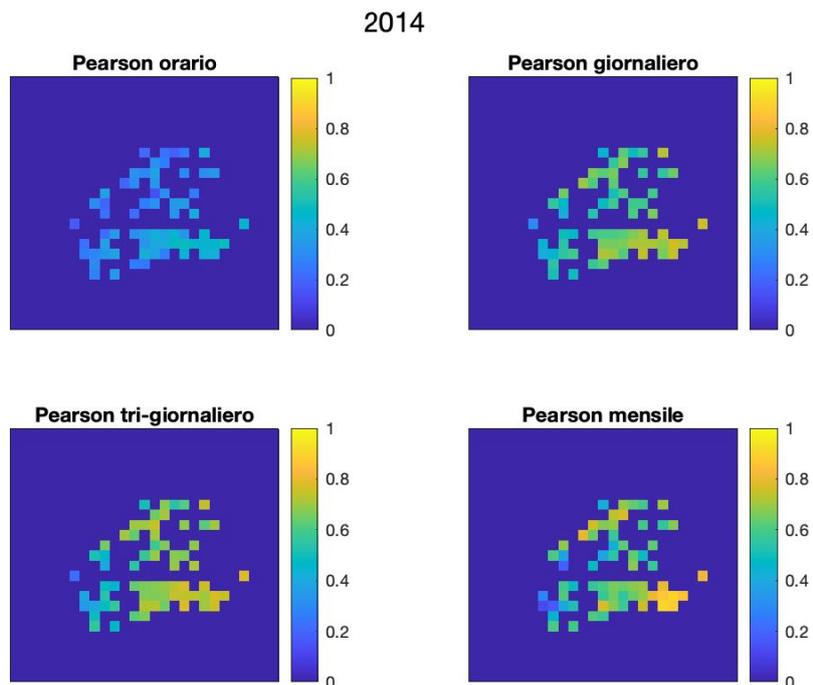


Figura 50 Distribuzione spaziale dell'indice di Pearson 2013: confrontando la distribuzione con quella dell'anno precedente si sottolinea un generale miglioramento, con valori nell'area collinare a sud-est prossimi all'unità. Questo significa che vi è buona correlazione tra le serie temporali di osservazione e stima.

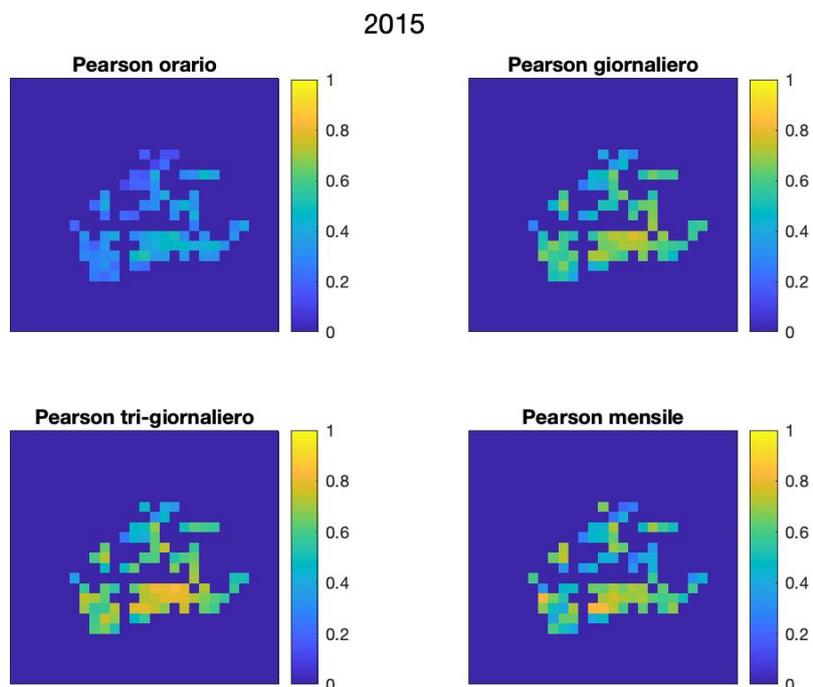


Figura 51 Distribuzione spaziale dell'indice di Pearson 2015: si noti il miglioramento dell'indice fino all'aggregazione tri-giornaliera, presentando poi valor inferiori nell'aggregazione mensile. Tuttavia, i valori dell'indice di Pearson sono elevati in tutta l'area meridionale del Piemonte, presentando ancora un'eccezione nelle celle a nord-ovest prossime ai rilievi alpini.

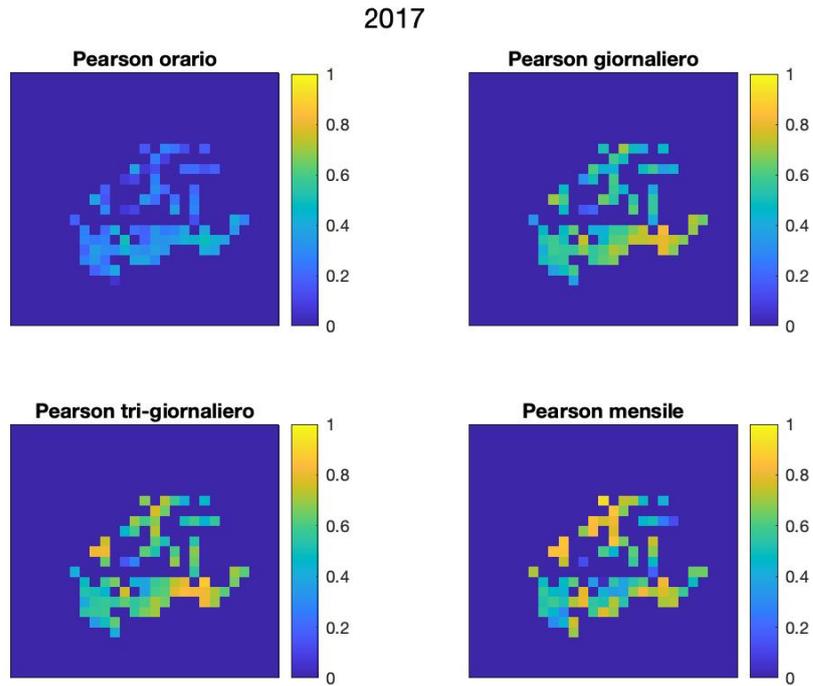


Figura 52 Distribuzione spaziale dell'indice di Pearson 2017: si osservi il migliorare dei valori dell'indice all'aumentare dell'aggregazione: vaste aree nelle aggregazioni dai 3 giorni in su si presentano superiori a 0,8. In questo anno, le aree a prestazioni minori sono quelle site a sud ovest, anch'esse prossime all'arco alpino.

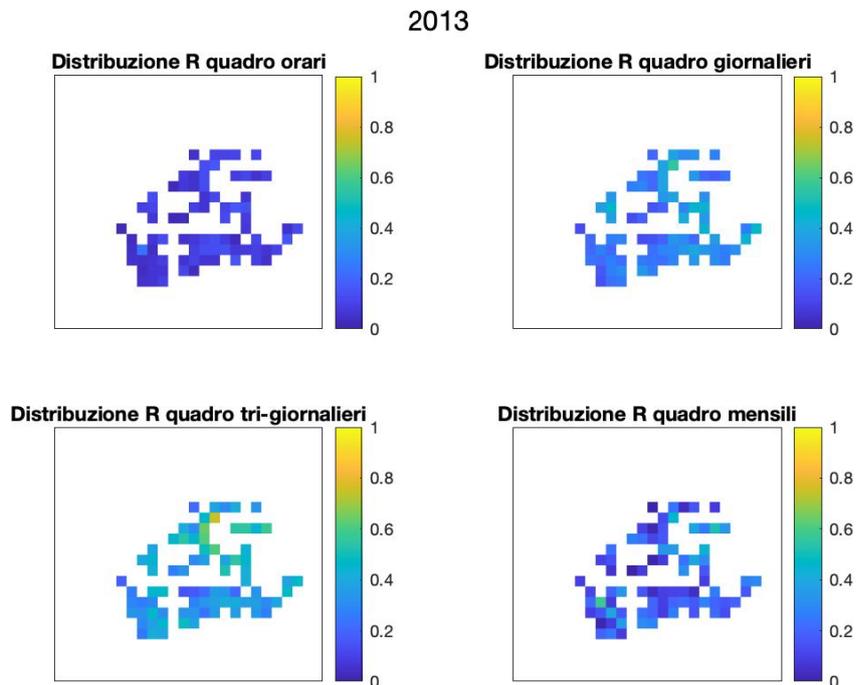


Figura 53 Distribuzione spaziale del coefficiente di determinazione 2013: vi è un aumento dei valori di R^2 con l'aumento dell'aggregazione temporale nei giornalieri e tri giornalieri, per poi subire una ricaduta nell'aggregazione mensile. In generale il coefficiente si mantiene inferiore a 0,6 in tutta la distribuzione.

2014

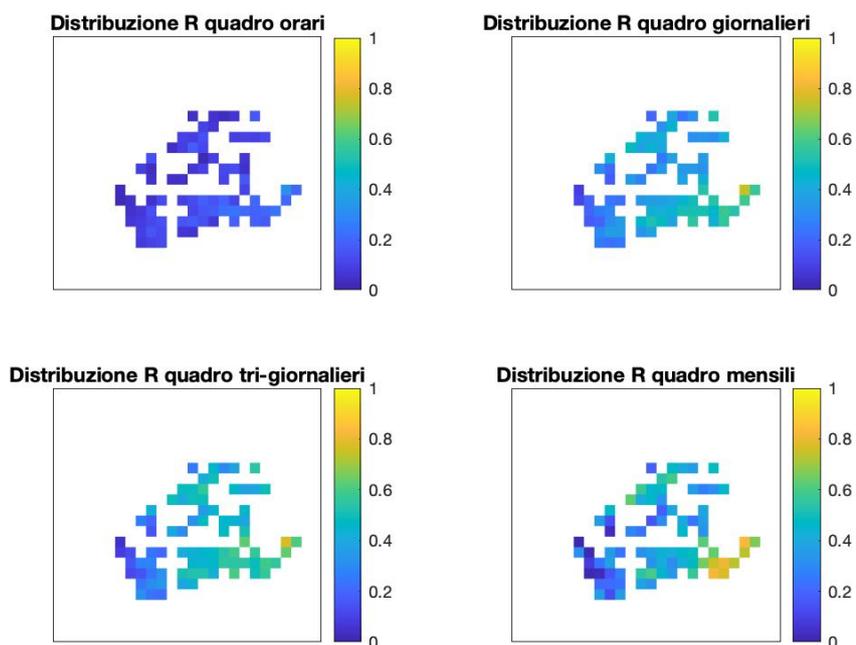


Figura 54 Distribuzione spaziale del coefficiente di determinazione 2014: si noti il costante miglioramento dell'indice in contemporanea all'aumento del periodo di aggregazione. Nell'area collinare a sud est, generalmente a prestazioni negative negli altri indici, il coefficiente R^2 presenta valori positivi prossimi all'unità.

2016

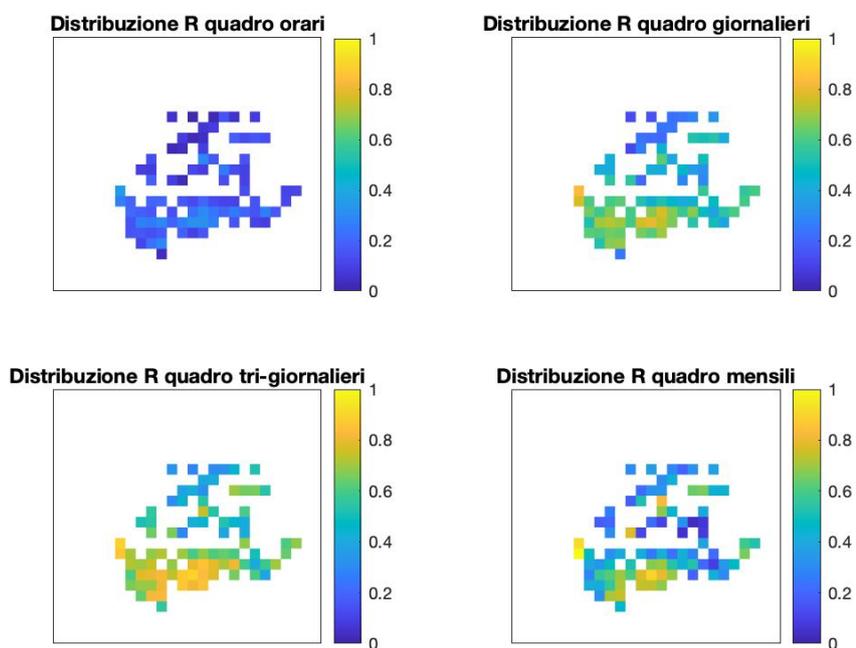


Figura 55 Distribuzione spaziale del coefficiente di determinazione 2016: si notino valori positivi e prossimi a 1 in tutta l'area meridionale della distribuzione. Tuttavia, nell'aggregazione mensile vi è una leggera riduzione di R^2 . Le aree in cui il coefficiente si presenta con valori ridotti si trovano a nord ovest ed a sud est, rispettivamente in prossimità dell'arco alpino e in zona collinare.

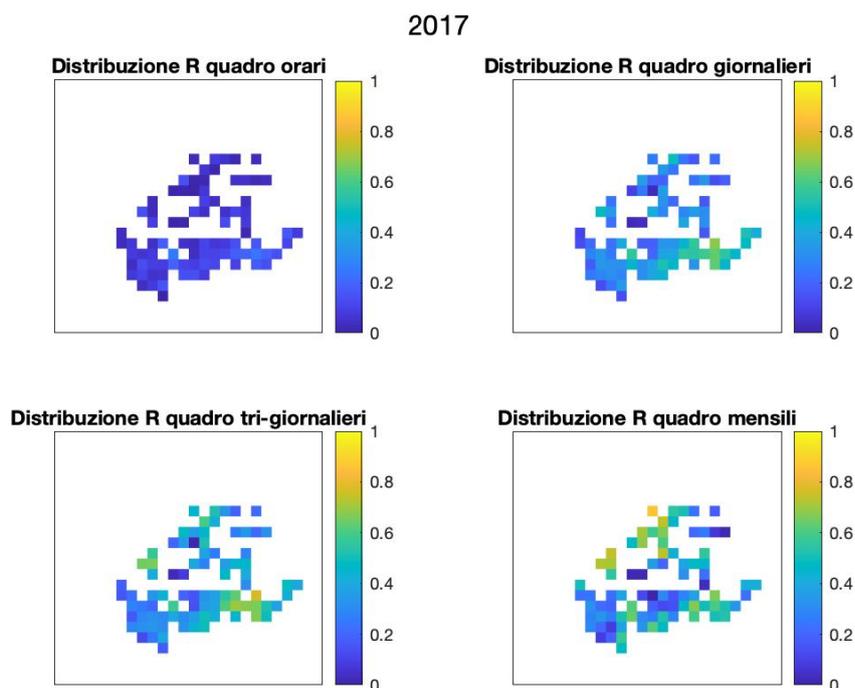
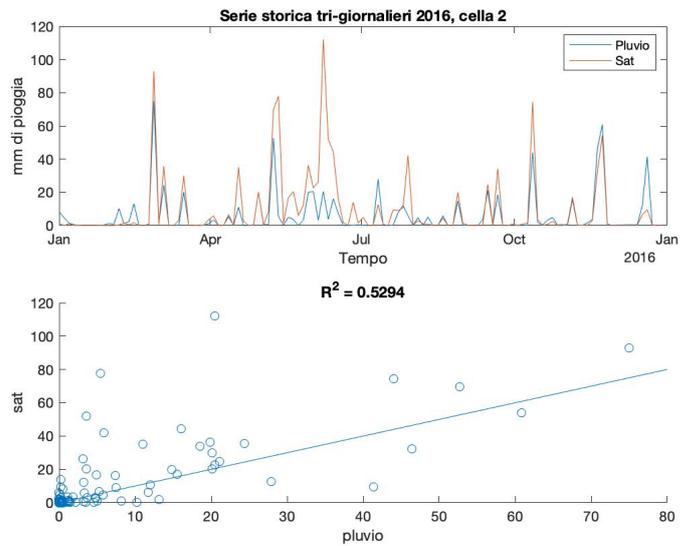
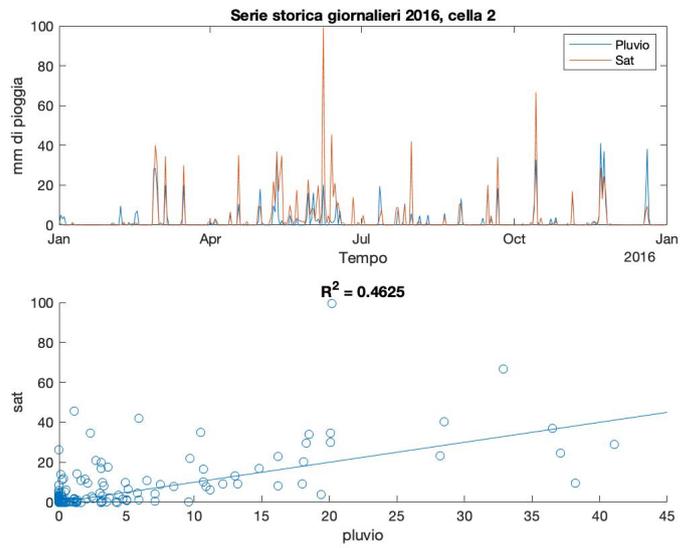
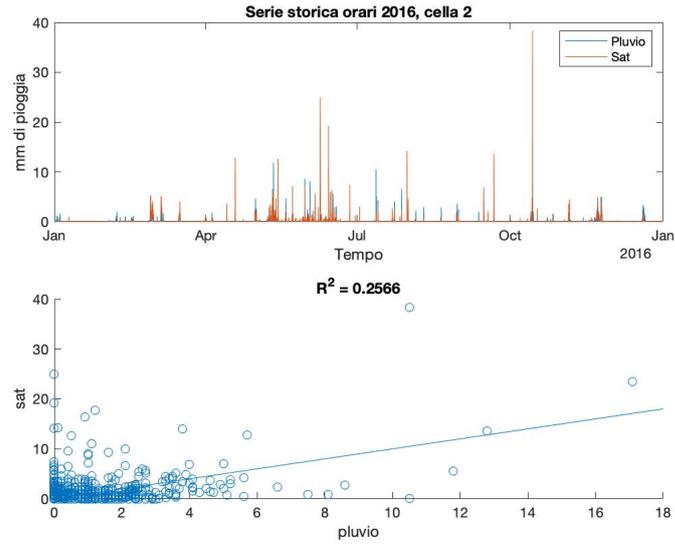


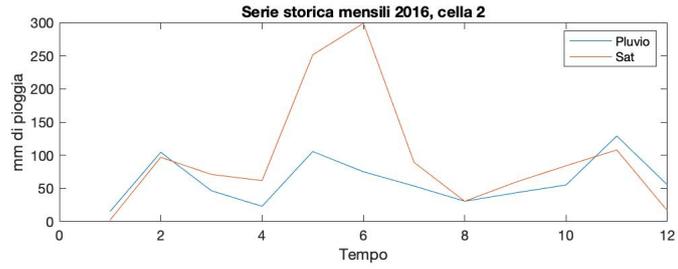
Figura 56 Distribuzione spaziale del coefficiente di determinazione 2017: rispetto agli anni precedenti i valori di R^2 appaiono nettamente inferiori in tutte le aggregazioni, pur presentando numerose celle in cui il coefficiente è superiore a 0,6.

Si riportano, inoltre, le rappresentazioni delle serie storiche e gli scatter plot prodotti per differenti aggregazioni temporali, riferiti alle celle numerate in Figura 29 non riportate nel testo principale. In tutte le aree considerate è possibile notare il miglioramento del coefficiente di determinazione R^2 nel passaggio tra aggregazione oraria e giornaliera. Tuttavia, i valori di R^2 maggiori si riscontrano nell'aggregazione di 3 giorni, eccezion fatta per la cella numero 3. Si noti in particolare, nell'aggregazione tri-giornaliera, la capacità della stima satellitare di individuare correttamente i giorni e i periodi di precipitazione, nonostante il quantitativo di pioggia stimata spesso non risulti esatto, causando un valore del coefficiente ridotto. Nell'area della griglia denominata cella 2, si osservi il coefficiente R^2 pari a 0,53, dovuto ad una generale sottostima da parte di IMERG. Nella cella numero 3, invece, si osservi il coefficiente molto elevato, pari a 0,89 nell'aggregazione di 3 giorni. Tale valore è da attribuirsi alla ridotta quantità di pioggia caduta a terra nell'area. Infine, la cella 5 risulta essere quella con maggiori differenze tra precipitazione osservata e stimata, con un R^2 massimo di 0,39. Osservando le serie storiche è possibile notare, inoltre, numerosi eventi la cui stima è pari a 0 mm a fronte, invece, di una reale precipitazione. Infatti, questa cella si trova nell'area collinare a sud-est in cui le prestazioni di stima sono risultate inferiori in tutti gli indici considerati, fornendo ulteriore conferma delle ipotesi di relazione tra prestazioni di stima ed elevazione media della cella.

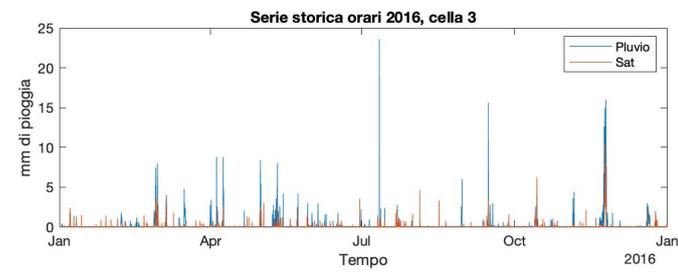
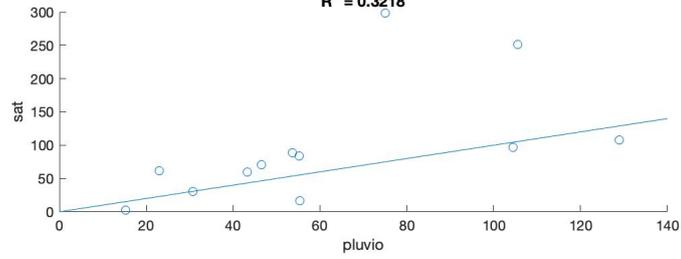
Appendice



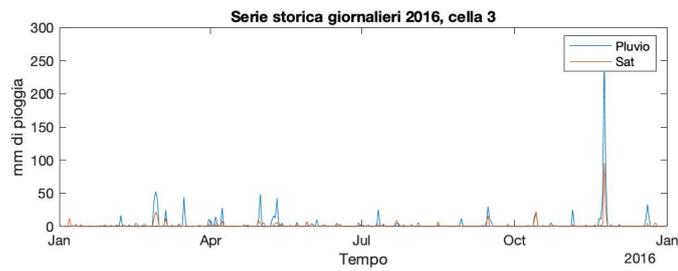
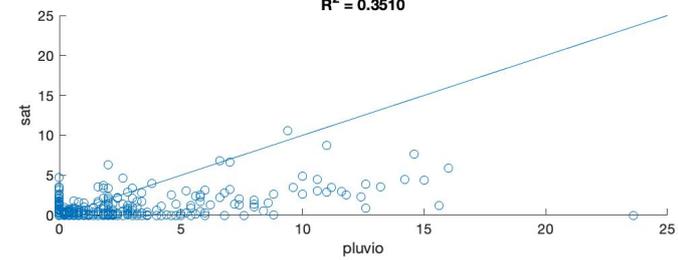
Appendice



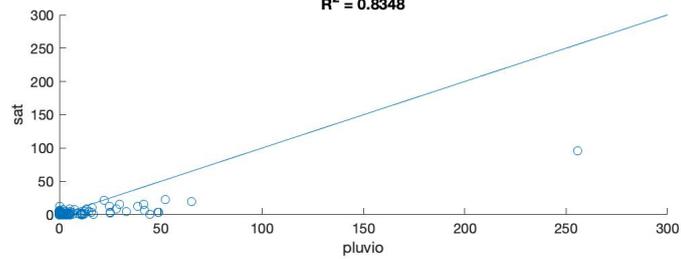
$R^2 = 0.3218$



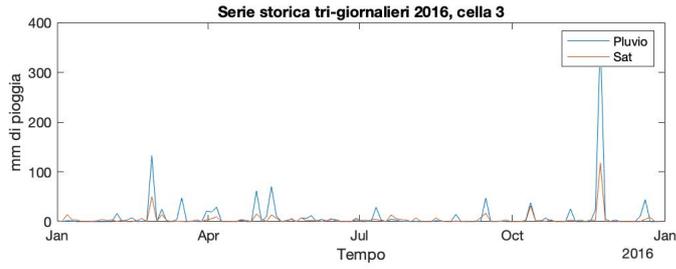
$R^2 = 0.3510$



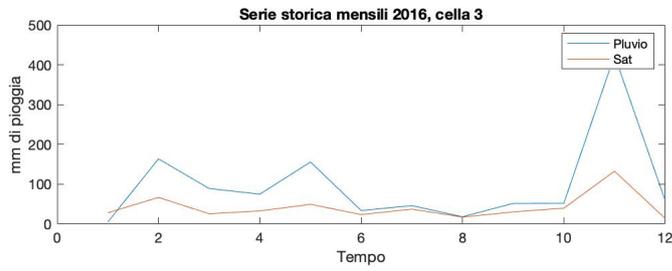
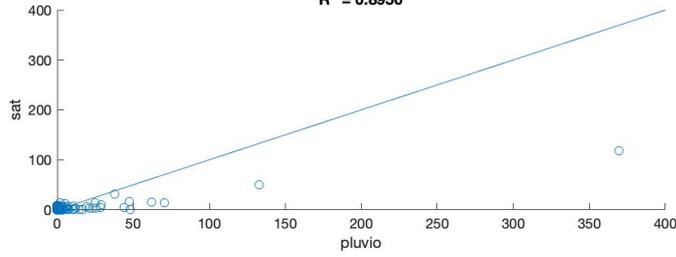
$R^2 = 0.8348$



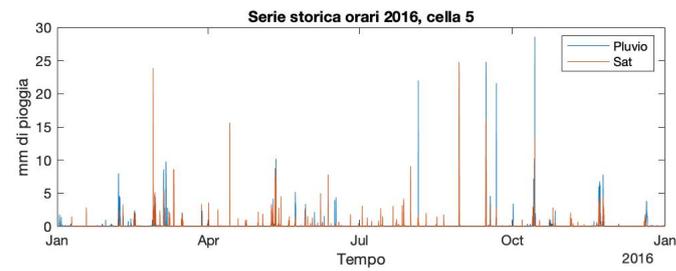
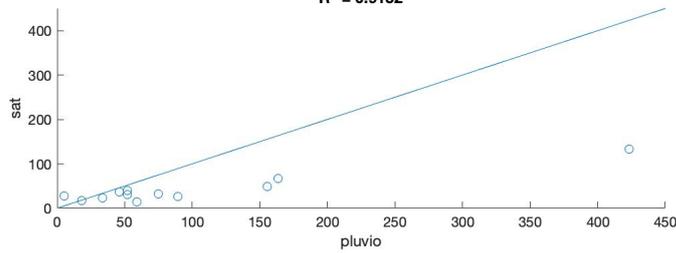
Appendice



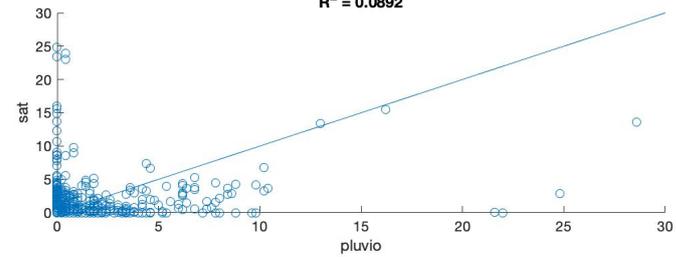
$R^2 = 0.8950$



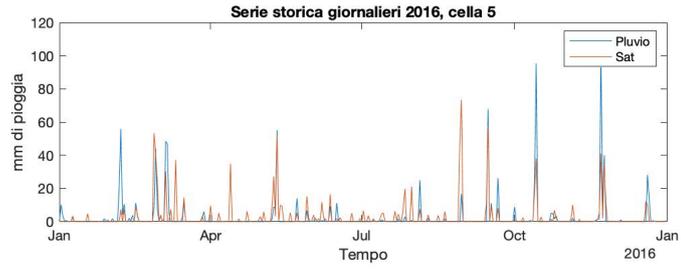
$R^2 = 0.9152$



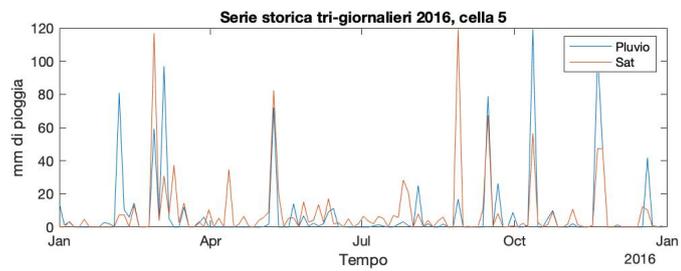
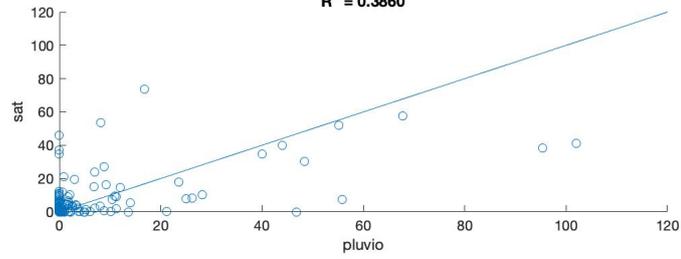
$R^2 = 0.0892$



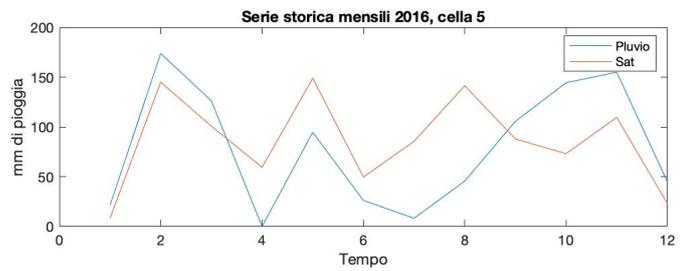
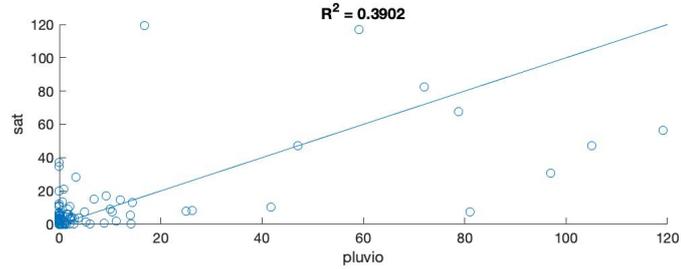
Appendice



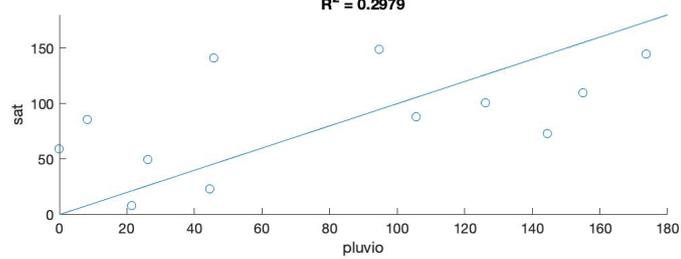
$R^2 = 0.3860$



$R^2 = 0.3902$



$R^2 = 0.2979$



Pre-processing script

```

close all
cd 2013
%% CALCOLO ACCUMULO PG DOVE ASSENTE
%devo importare il file 3a e chiamarlo "pluviodati"
minutiNUM=str2double(pluviodati.Minuti);
pioggia=pluviodati.Pioggia;
pioggiaORIGINALE=pioggia;
k=1;
for i=1:size(pluviodati,1)
    if minutiNUM(i)==10
        pioggia(i+5)=sum(pioggia(i:i+5));
        idsum(k)=i;
        k=k+1;
    end
end
%% then check
diversi=pioggiaORIGINALE- pioggia;
dove_nan=isnan(diversi); %ci sono NaN?
diversi(dove_nan)=0;
dovediversi= diversi~=0; %qui qualcosa e' cambiato, escludendo i nan
quantodiversi=nnz(diversi)

pluviodati.Pioggia(dovediversi)=pioggia(dovediversi);
stazMINelim=pluviodati.Stazione(idsum);
stazMINelim=unique(stazMINelim,'stable');
%% elimino i dati NON ORARI
%cerco i valori di minuti
minuti_da_eliminare=minutiNUM==10|minutiNUM==20|minutiNUM==30|minutiNUM=
=40|minutiNUM==50;
pluviodati_minuti_eliminati=pluviodati(~minuti_da_eliminare,:);
clear diversi dove_nan dovediversi i idsum k pioggia quantodiversi
%% pulire pluviodati_fase 1
pluviodatiORIGINAL=pluviodati; %i pluviodati originali
pluviodati=pluviodati_minuti_eliminati;
a1=pluviodati(:,1); %la prima colonna e' IDmeteo
a1=table2array(a1);
alnan=isnan(a1);
a2=pluviodati(:,2); %seconda colonna sono stazioni
a2=table2array(a2);
CHECK=a2(alnan) %controllo che le stazioni non esistano: se va tutto
bene
%dovrei avere un empty array
%% pulire pluviodati_fase 2
pluviodati(alnan,:)=[];

a6=pluviodati(:,6); %pioggia
a6=table2array(a6);
a6nan=isnan(a6);
pluviodati(a6nan,:)=[];
%controllo che tutti i minuti siano NaN o 0
a5=pluviodati(:,5);
a5=table2array(a5);
a5=str2double(a5);

```

```

a5nan=isnan(a5);
a5uguale0 = a5==0;
a5nan=a5nan|a5uguale0;
%I MINUTI DEVONO ESSERE =0 OPPURE NAN
pluviodati(~a5nan,:)=[];
clear a1 a2 a5 a6 a7 all alnan a5nan a6nan a7nan allnan a5uguale0 check
%% Assegnazione ID alle stazioni 3a
%NB la LAT decresce, la LON cresce
%includo la lat_anag e lon_anag nella stessa griglia di lat e lon,
%attraverso gli indici di lat e lon
%LATITUDINE
k=1;
for j=1:size(lat_agg,1)
    for i=2:length(lat)
        if (lat_agg(j)<lat(i-1) && lat_agg(j)>lat(i))
            %uso i e i-1, con i che parte da 2, cosi evito il problema
del
            %superare l'indice i all'interno del vettore lat
            idx_lat_agg(k)=i-1;
            k=k+1;
        end
    end
end
clear i j

%LONGITUDINE
k=1;
for j=1:size(lon_agg,1)
    for i=2:(length(lon)-1)
        if (lon_agg(j)>lon(i-1) && lon_agg(j)<lon(i))
            %cambiano i segni perche' la lon e' crescente
            idx_lon_agg(k)=i-1;
            k=k+1;
        end
    end
end
clear i j k
%prendo la lon/lat in corrispondenza dell'indice e aggiungo 0.05 per
avere
%le coordinate nel centro del quadrato, di maglia 0.1 gradi
dimensioneLon_Lat=length(lon_agg);
lon_agg_grid = zeros(dimensioneLon_Lat,1);
lat_agg_grid = zeros(dimensioneLon_Lat,1);
for i=1:length(idx_lon_agg)
    lon_agg_grid(i)=lon(idx_lon_agg(i))+0.05;
    lat_agg_grid(i)=lat(idx_lat_agg(i))+0.05;
end
%% assegno coord ai dati pluvio
stazione_pluvio1=pluviodati.Stazione;
lat_pluvio=NaN.*zeros(height(pluviodati),1);
lon_pluvio=NaN.*zeros(height(pluviodati),1);
for i=1:length(lat_pluvio)
    for j=1:size(stazione_agg,1)
        if isequal(stazione_pluvio1{i}, stazione_agg{j})

```

```

        lat_pluvio(i)=lat_agg_grid(j);
        lon_pluvio(i)=lon_agg_grid(j);
    end
end
end
%% check
%in caso ci siano errori ci saranno dei NaN. In quali stazioni?
doveNaN=isnan(lat_pluvio);
doveNaN=pluviodati.Stazione(doveNaN);
doveNaN=unique(doveNaN,'stable')
%SOLUZIONE=AGGIUNGO e poi ripeto sezione
%% aggiungo
stazione_aggWORK=stazione_agg;
ultima_riga=length(stazione_agg);
x=length(doveNaN);
stazione_aggWORK(ultima_riga+1:ultima_riga+x)=doveNaN;
stazione_aggWORK=stazione_aggWORK([1:11 ultima_riga+1 12:61
ultima_riga+2 ...
    62:64 ultima_riga+3 65:87 ultima_riga+4 88:90 ...
    ultima_riga+5 91:95 ultima_riga+6 96:100 ultima_riga+7 101:119 ...
    ultima_riga+8 120:131 ultima_riga+9 132:ultima_riga]);
stazione_agg=stazione_aggWORK;
clear stazione_aggWORK
%% aggiungo coordinate
%devo aggiungere anche le coordinate
%latitudine
lat_agg(ultima_riga+1)=44.618471;
lat_agg(ultima_riga+2)=44.46517;
lat_agg(ultima_riga+3)=44.459278;
lat_agg(ultima_riga+4)=44.67821;
lat_agg(ultima_riga+5)=44.45814;
lat_agg(ultima_riga+6)=44.586317;
lat_agg(ultima_riga+7)=44.758067;
lat_agg(ultima_riga+8)=44.821416;
lat_agg(ultima_riga+9)=44.690298;
lat_agg=lat_agg([1:11 ultima_riga+1 12:61 ultima_riga+2 ...
    62:64 ultima_riga+3 65:87 ultima_riga+4 88:90 ...
    ultima_riga+5 91:95 ultima_riga+6 96:100 ultima_riga+7 101:119 ...
    ultima_riga+8 120:131 ultima_riga+9 132:ultima_riga]);
%longitudine
lon_agg(ultima_riga+1)=7.949471;
lon_agg(ultima_riga+2)=7.547300;
lon_agg(ultima_riga+3)=7.433389;
lon_agg(ultima_riga+4)=8.14579;
lon_agg(ultima_riga+5)=8.10081;
lon_agg(ultima_riga+6)=7.953544;
lon_agg(ultima_riga+7)=8.398557;
lon_agg(ultima_riga+8)=8.128007;
lon_agg(ultima_riga+9)=8.066328;
lon_agg=lon_agg([1:11 ultima_riga+1 12:61 ultima_riga+2 ...
    62:64 ultima_riga+3 65:87 ultima_riga+4 88:90 ...
    ultima_riga+5 91:95 ultima_riga+6 96:100 ultima_riga+7 101:119 ...
    ultima_riga+8 120:131 ultima_riga+9 132:ultima_riga]);
clear stazione_aggWORK ultima_riga
%poi devo tornare ad ASSEGNAZIONE ID STAZIONI 3A

```

```

%% unisco colonne
pluviodati.lon=lon_pluvio;
pluviodati.lat=lat_pluvio;
%% assegno indice ai pluviodati
% Per creare una matrice sovrapponibile
% non avendo una mappa e' meglio usare gli indici da 1 a 25 o 27 come
% nei dati nasa, e lasciare le coordinate su un array separato

%LATITUDINE
lat_idx=NaN.*zeros(height(pluviodati),1);
k=1;
for j=1:size(lat_pluvio,1)
    for i=2:length(lat)
        if (lat_pluvio(j)<lat(i-1) & lat_pluvio(j)>lat(i))
            %uso i e i-1, con i che parte da 2, cosi evito il problema
del
            %superare l'indice i all'interno del vettore lat
            lat_idx(k)=i-1;
            k=k+1;
        end
    end
end

%LONGITUDINE
lon_idx=NaN.*zeros(height(pluviodati),1);
k=1;
for j=1:size(lon_pluvio,1)
    for i=2:(length(lon)-1)
        if (lon_pluvio(j)>lon(i-1) & lon_pluvio(j)<lon(i))
            %cambiano i segni perche' la lon e' crescente
            lon_idx(k)=i-1;
            k=k+1;
        end
    end
end
clear i j k
%% unisco colonne indice
pluviodati.lon_idx=lon_idx;
pluviodati.lat_idx=lat_idx;
%% aggiusto data/ora nel file pluviodati
data=datetime(pluviodati.Data);
ora=hours(pluviodati.Ora)-hours(1);
data=data + ora;
pluviodati.Data = data;
%% Trasformo in matrice i dati pluvio
k=1; %indichera' il "piano" della matrice
pluviomatrix=NaN.*zeros(25,27,(365*24)); %da cambiare in caso bisestile
anno=year(pluviodati.Data(1)); %ho 1 file per anno, basta 1
data_indagine=datetime(anno,1,1,0,0,0);
un_ora=hours(1);
pioggia=pluviodati(:,6);
pioggia=table2array(pioggia);
%inserisco nella pluviomatrix
for k=1:(365*24) %1 anno

```

```

ciclodata= pluviodati.Data==data_indagine;
for n=1:27 %lon. Guardo colonna x colonna
    for m=1:25 %lat. Scorro riga
        pluviomatrix(m,n,k)= ...
            mean(pioggia(ciclodata & lon_idx==n & lat_idx==m));
    end
end
data_indagine=data_indagine+un_ora;
disp(k)
end
%% Creo un indice di date dei "piani" della matrice + creo tabella
riepilogo
piano_matrice=pluviodati.Data;
piano_matrice=unique(piano_matrice);
Mese=(1:12)';

```

Analisi statistica

Sono state omesse alcune parti relative alla produzione di immagini o tabelle secondarie

```

close all
cd 2013
%% creo matrice dati satellite
pluvio=pluviomatrix;
cd (...)
load('risultati_piemonte_201301.mat')
precipitation_cal2013=precipitation_cal; %gennaio
for i=2:12
    switch i %apro il mese nasa
        case 2
            load('risultati_piemonte_201302.mat')
        case 3
            load('risultati_piemonte_201303.mat')
        case 4
            load('risultati_piemonte_201304.mat')
        case 5
            load('risultati_piemonte_201305.mat')
        case 6
            load('risultati_piemonte_201306.mat')
        case 7
            load('risultati_piemonte_201307.mat')
        case 8
            load('risultati_piemonte_201308.mat')
        case 9
            load('risultati_piemonte_201309.mat')
        case 10
            load('risultati_piemonte_201310.mat')
        case 11
            load('risultati_piemonte_201311.mat')
    end
end

```

```

        case 12
            load('risultati_piemonte_201312.mat')
        end
        precipitation_cal2013=cat(3,precipitation_cal2013,
precipitation_cal);
    end
    cd (...)
    save('precipitation_cal2013.mat','precipitation_cal2013')
    %% rendo orario precipitation_cal2013
    pari=2:2:size(precipitation_cal2013,3);
    dispari=1:2:size(precipitation_cal2013,3);
    satdatiYearYear=precipitation_cal2013;
    satdatiYear(:, :, dispari)=precipitation_cal2013(:, :, dispari)+precipitatio
n_cal2013(:, :, pari);
    satdatiYear1=satdatiYear./2;
    sat=satdatiYear1(:, :, dispari);
    clear satdatiYear satdatiYearYear pari dispari
    %% tabella contingenza oraria
    % CODICE
    % 1=CORRECT NEGATIVES
    % 2=MISSES
    % 3=FALSE ALARM
    % 4=HIT
    correctNegH=0;
    missesH=0;
    falseAlarmsH=0;
    hitsH=0;
    contingencyH=NaN.*pluvio;
    RMSEh=NaN.*zeros(25,27);
    PmenoS=NaN.*pluvio; %serve per l'RMSE
    A=NaN.*pluvio; %numeratore Pearson
    B=NaN.*pluvio; %primo elemento denominatore Pearson
    C=NaN.*pluvio; %secondo elemento denominatore Pearson
    rH=NaN.*zeros(25,27); %Pearson correlation coeff orario
    % The r measures how well the estimated values correspond to the
    observed
    % values. The RMSE is a commonly used measure of the differences between
    % the observed precipitation and the estimated precipitation
    sogliaH=0.2; %threshold di pioggia (considero PG tutto quello maggiore)
    for n=1:27 %lon. Guardo colonna x colonna
        for m=1:25 %lat. Scorro riga
            Pmean=mean(pluvio(m,n,:));
            Smean=mean(sat(m,n,:));
            for k=1:size(pluvio,3)
                P=pluvio(m,n,k);
                S=sat(m,n,k);
                if S<sogliaH & P<sogliaH
                    contingencyH(m,n,k)=1;
                    correctNegH=correctNegH+1;
                elseif S<sogliaH & P>=sogliaH
                    contingencyH(m,n,k)=2;
                    missesH=missesH+1;
                elseif S>=sogliaH & P<sogliaH
                    contingencyH(m,n,k)=3;
                    falseAlarmsH=falseAlarmsH+1;
                end
            end
        end
    end

```

```

elseif S>=sogliaH & P>=sogliaH
    contingencyH(m,n,k)=4;
    hitsH=hitsH+1;
end
PmenoS(m,n,k)=P-S;
A(m,n,k)=(P-Pmean).*(S-Smean);
B(m,n,k)=(P-Pmean).^2;
C(m,n,k)=(S-Smean).^2;
end
end
end
for n=1:27
    for m=1:25
        RMSEh(m,n)=sqrt((1/size(pluvio,3)).*sum((PmenoS(m,n,:)).^2));

rH(m,n)=(sum(A(m,n,:)))/(sqrt(sum(B(m,n,:))).*sqrt(sum(C(m,n,:))));
end
end
%calcolo percentuali
totH=correctNegH+missesH+falseAlarmsH+hitsH;
correctNegHx100=correctNegH/totH;
missesHx100=missesH/totH;
falseAlarmsHx100=falseAlarmsH/totH;
hitsHx100=hitsH/totH;
RMSEh_max=max(max(RMSEh));
clear A B C PmenoS Pmean Smean
figure(1)
imagesc(RMSEh)
title('RMSE orario')
colorbar
saveas(gcf,'RMSEh.jpg')
figure(2)
imagesc(rH)
title('r orario')
colorbar
saveas(gcf,'rH.jpg')
%% risultati orari_ INDICI CATEGORICI
% FAR:False Alarm Ratio
% POD:Probability of Detection
% CSI:Critical Success Index
FARh=falseAlarmsH/(hitsH+falseAlarmsH); %ottimale se vicino 0
PODh=hitsH/(hitsH+missesH); %ottimale se vicino 1
CSIh=hitsH/(hitsH+falseAlarmsH+missesH); %ottimale se vicino 1
BIASH=(hitsH+falseAlarmsH)/(hitsH+missesH); %ottimale se vicino 1
%% tabella contingenza giornaliera
[y,m,d]=ymd(piano_matrice);
data_aggregazioni=datetime(y,m,d);
clear y m d

data_sequenziale=unique(data_aggregazioni);
dayPL13=NaN.*zeros(25,27,length(data_sequenziale)); %OCCHIO AI BISESTILI
daySAT13=NaN.*zeros(25,27,length(data_sequenziale));
for i=1:length(data_sequenziale) %tutto2013
    idx=data_aggregazioni==data_sequenziale(i);

```

```

for n=1:27 %lon. Guardo colonna x colonna
    for m=1:25
        dayPL13(m,n,i)=sum(pluvio(m,n,idx));
        daySAT13(m,n,i)=sum(sat(m,n,idx));
%Usa SUM e non NANSUM in modo da mantenere i NaN. Inoltre in questo
modo,
%le caselle con un solo valore nan negli orari saranno in automatico
poste a
%nan. E' un'ulteriore pulizia. Inoltre, se provassi a usare nansum, il
%numero di CORRECTNEG schizzerebbe su, perch? tutti i nan vengono
%sostituiti con 0. con i relativi indici categorici che assumerebbero
%valori non corretti.
    end
end
end
% CODICE
% 1=CORRECT NEGATIVES
% 2=MISSES
% 3=FALSE ALARM
% 4=HIT
correctNeg=0;
misses=0;
falseAlarms=0;
hits=0;
contingency=NaN.*dayPL13;
RMSE=NaN.*zeros(25,27);
PmenoS=NaN.*dayPL13; %serve per l'RMSE
A=NaN.*dayPL13; %numeratore Pearson
B=NaN.*dayPL13; %primo elemento denominatore Pearson
C=NaN.*dayPL13; %secondo elemento denominatore Pearson
r=NaN.*zeros(25,27); %Pearson orario
soglia=0.2; %threshold
    for n=1:27 %lon. Guardo colonna x colonna
        for m=1:25 %lat. Scorro riga
            Pmean=mean(dayPL13(m,n,:));
            Smean=mean(daySAT13(m,n,:));
            for k=1:size(dayPL13,3)
                P=dayPL13(m,n,k);
                S=daySAT13(m,n,k);
                if S<soglia & P<soglia
                    contingency(m,n,k)=1;
                    correctNeg=correctNeg+1;
                elseif S<soglia & P>=soglia
                    contingency(m,n,k)=2;
                    misses=misses+1;
                elseif S>=soglia & P<soglia
                    contingency(m,n,k)=3;
                    falseAlarms=falseAlarms+1;
                elseif S>=soglia & P>=soglia
                    contingency(m,n,k)=4;
                    hits=hits+1;
                end
                PmenoS(m,n,k)=P-S;
                A(m,n,k)=(P-Pmean).*(S-Smean);
                B(m,n,k)=(P-Pmean).^2;
            end
        end
    end

```

```

        C(m,n,k)=(S-Smean).^2;
    end
end
end
for n=1:27
    for m=1:25
        RMSE(m,n)=sqrt((1/size(dayPL13,3)).*sum((PmenoS(m,n,:)).^2));

r(m,n)=(sum(A(m,n,:)))/(sqrt(sum(B(m,n,:))).*sqrt(sum(C(m,n,:))));
        end
end
%% calcolo percentuali --> uguale agli orari
%% risultati --> uguale agli orari
%% ripeto per altre aggregazioni
%% SCATTER E CORRELAZIONI
PLesiste=~isnan(dayPL13(:,:,1));
PLquantoEsiste=0;
PLcoordEsiste=zeros(80,2);
for n=1:27
    for m=1:25
        if PLesiste(m,n)==1
            PLquantoEsiste=PLquantoEsiste+1;
            PLcoordEsiste(PLquantoEsiste,1)=m;
            PLcoordEsiste(PLquantoEsiste,2)=n;
        end
    end
end
end
close all

%R quadri ORARI
RsquaredH=[PLcoordEsiste (NaN.*zeros(PLquantoEsiste,1))];
%tabella con coordinate cella e valore di R quadro
RsquaredMapH=NaN.*PLesiste;
for i=1:PLquantoEsiste
    m=PLcoordEsiste(i,1);
    n=PLcoordEsiste(i,2);
    P=squeeze(pluvio(m,n,:));
    S=squeeze(sat(m,n,:));
    R=corrcoef(P,S,'rows','complete'); %omette tutte le
righe(=valori)con NAN
    R_squared=R(1,2)^2;
    RsquaredH(i,3)=R_squared;
    RsquaredMapH(m,n)=R_squared;
end
figure(10)
imagesc(RsquaredMapH)
title('Distribuzione R quadro orari')
colorbar
% R quadri giornalieri
RsquaredDaily=[PLcoordEsiste (NaN.*zeros(PLquantoEsiste,1))];
RsquaredMapDaily=NaN.*PLesiste;
for i=1:PLquantoEsiste
    m=PLcoordEsiste(i,1);
    n=PLcoordEsiste(i,2);

```

```

P=squeeze(dayPL13(m,n,:));
S=squeeze(daySAT13(m,n,:));
R=corrcoef(P,S,'rows','complete');
R_squared=R(1,2)^2;
RsquaredDaily(i,3)=R_squared;
RsquaredMapDaily(m,n)=R_squared;

%    %possibilita' di plottare gli SCATTER (1 x ogni cella=86 scatter)

% %    celle_campione=1:10:PLquantoEsiste;
% %    if any(i==celle_campione)
%    xMax=max(max(P),max(S));
%    figure(i+20)
%    scatter(P,S)
%    xlabel('pluvio')
%    ylabel('sat')
%    xlim([1 xMax])
%    ylim([1 xMax])
%    reline(1,0)
%    title(sprintf('cella (%u,%u), R^2 = %.4f',m,n,R_squared))
%%possibilita' di plottare SERIE STORICA
%    figure(i+PLquantoEsiste+20)
%    plot(data_sequenziale,P)
%    hold on
%    plot(data_sequenziale,S)
%    xlabel('Giorni')
%    ylabel('mm di pioggia')
%    title(sprintf('Serie storica 2013, casella (%u,%u)',m,n))
%    legend('Pluvio','Sat')
%    hold off
% %    end
end

figure(11)
imagesc(RsquaredMapDaily)
title('Distribuzione R quadro giornalieri')
colorbar
%% ripeto per altre aggregazioni
%% creo tabella riepilogo prestazioni
%% grafico RMSElim
RMSElimPearson13(RMSEh, RMSE, RMSE3, RMSEm, rH, r, r3, rM)
%% grafici R quadro
pcolorRquadroSerieStoriche13(RsquaredMapH, RsquaredMapDaily, RsquaredMap3,
RsquaredMapM, lon, lat)
%% PLOTTO serie + scatter
SerieScatter13(pluvio, sat, piano_matrice)
SerieScatter13(dayPL13, daySAT13, data_sequenziale)
data3=data_sequenziale(1:3:end);
SerieScatter13(day3PL13, day3SAT13, data3)
mesi=1:12;
SerieScatter13(monPL13, monSAT13, mesi)
close all

cd (...)

```

Appendice

```
save('Rsq13.mat','RsquaredH','RsquaredDaily','Rsquared3','RsquaredM')
%% tabella RMSE
tabellaRMSE13(RMSEh, RMSE, RMSE3, RMSEm)
%% correctnegatives+hits
ZZZcorrectnegPIUhits(correctNegHx100, correctNegx100, correctNeg3x100, ...
    correctNegMx100, hitsHx100, hitsx100, hits3x100, hitsMx100)
```


Ringraziamenti

Giungere al termine di questo percorso di studi è una grande soddisfazione e, se sono arrivato a completarlo, è merito delle numerose persone che mi sono state a fianco e hanno contribuito, ognuno a suo modo, a rendere questo periodo della mia vita un periodo felice.

Desidero ringraziare i componenti della 3a, Massimo, Emanuela, Laura e Gianmaurizio, per avermi accolto dandomi l'opportunità di partire dal frutto del loro lavoro per costruire questa tesi.

Un grazie speciale va a Ricky, Bea e Davide, che hanno saputo trasmettere anche a distanza il loro affetto e le loro competenze tecniche.

Alla mia famiglia, Carolina, Laura, Ezio, Ricky e le mie nonne, rivolgo i ringraziamenti più sentiti per avermi sostenuto in tutte le fasi della vita universitaria, con preziosi consigli e con la loro esperienza di cui io non posso ancora vantarmi.

Infine, un grazie immenso ai numerosi amici e compagni, a quelli incontrati sui banchi e quelli che ci sono da sempre, senza i quali questi anni non avrebbero avuto la luce e la spensieratezza necessarie.