



Interconfronto SWE

Pian de L'Eyvia - Pila
5 Marzo 2015

Report tecnico a cura di G. Filippa, E. Cremonese, F. Diotri, U. Morra di Cella - ARPA Valle d'Aosta

In sintesi

- 186 misure di SWE effettuate con tre metodi diversi su una superficie di circa 1000 m², 34 operatori nivologici coinvolti nell'esperimento (fig. 1).
- Una coppia di rilevatori che effettua due misure successive di SWE su un manto nevoso nelle stesse condizioni e con lo stesso metodo commette un errore medio del 3% (fig. 4).
- Un gruppo di 17 coppie di rilevatori che effettua misure lungo un transetto di circa 12-15 m lineari in condizioni di manto nevoso ritenuto omogeneo commette un errore medio del 7% (fig. 4). La variabilità interoperatore osservata (anche dell'ordine di 100 mm di SWE, fig. 5) sottolinea la necessità di ulteriori esercizi di interconfronto finalizzati ad uniformare la tecnica di campionamento e la realizzazione delle misure.
- La misura della densità media ponderata di un profilo di neve effettuata secondo strati omogenei (metodo AINEVA) porta ad una stima dell'equivalente in acqua del manto nevoso confrontabile con la misura condotta a profondità fisse (metodo SWE) (figg. 2 e 6).
- La misura dell'equivalente in acqua del manto nevoso effettuata con il tubo EV ha portato, in questo esperimento, ad una sottostima dell'SWE (figg. 2 e 6).

1 I dati

Numerosità campionaria

Hanno partecipato all'attività 17 coppie di misuratori e sono state complessivamente effettuate 186 misure. Di queste, 84 misure a profondità fisse (metodo di seguito chiamato SWE) e altrettante secondo gli strati (metodo AINEVA), mentre 18 misure sono state effettuate con il tubo EV. 76 di queste misure hanno avuto luogo nel plot A, 70 nel plot B e 40 nel plot C (Fig.1).

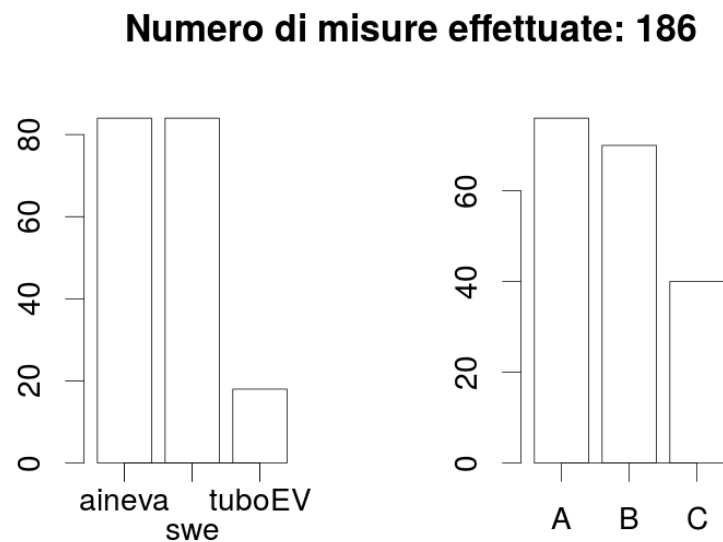


Figura 1: Numero di misure effettuate per ciascun metodo ed in ciascun plot

Medie per metodo e per plot

La figura 2 mostra i valori mediani di SWE, HS e RHO per ciascuno dei tre metodi. La mediana per un determinato metodo è quindi calcolata mediando le misure di tutte le coppie (incluse le ripetizioni) e di tutti i plot. La barra di errore rappresenta la mean absolute deviation (MAD), statistica non parametrica concettualmente analoga alla deviazione standard. Si riporta inoltre il coefficiente di variazione, calcolato come rapporto tra MAD e mediana $\times 100$. La linea rossa rappresenta la mediana generale. Per quanto riguarda SWE, osserviamo che le misure effettuate col tubo EV sono decisamente inferiori rispetto agli altri due metodi. La variabilità del dato (errore%) è invece confrontabile tra metodi. HS risulta sostanzialmente identica per i tre metodi e la sua variabilità è inferiore al 5%. Il dato di densità presenta caratteristiche analoghe a quelle descritte per SWE. Come detto, la linea rossa rappresenta la mediana generale. Non potendo ipotizzare a priori maggiore affidabilità di un metodo rispetto ad un altro, la mediana generale diverrà la nostra verità a terra, ossia il numero contro il quale sarà successivamente testata l'ipotesi se un dato metodo sovrastimi o sottostimi il reale valore di SWE.

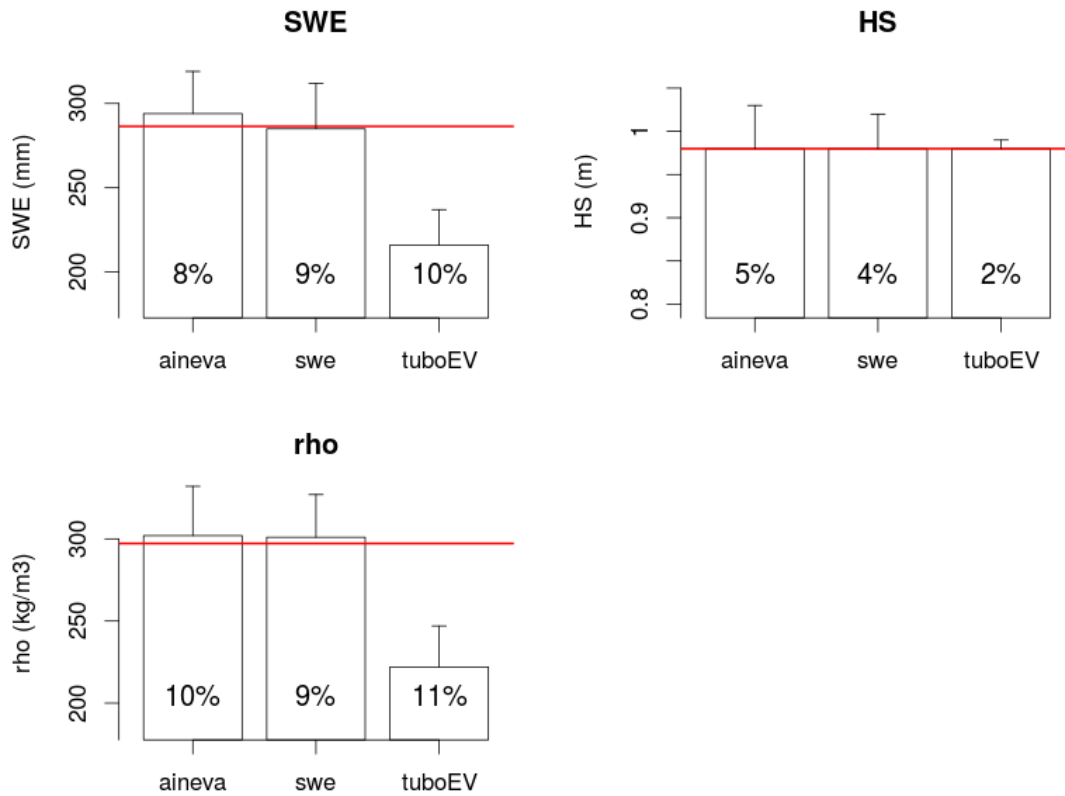


Figura 2: Valori mediani di SWE, HS e RHO in funzione del metodo utilizzato. La barra di errore rappresenta la mean absolute deviation (MAD). Si riporta inoltre il coefficiente di variazione, calcolato come $MAD/mediana * 100$. La linea rossa rappresenta la mediana generale.

La figura 3 è l'analogo di fig. 2 per la variabilità tra plot. Si osserva una altezza media del manto nevoso inferiore in C, cui però corrisponde una densità più elevata. I due comportamenti opposti di HS e RHO si compensano in qualche modo nel calcolo di SWE. Ciò fa sì che non vi sia variabilità spaziale per SWE. Tuttavia i coefficienti di variazione spaziale

sono progressivamente maggiori rispetto ai coefficienti di variazione intrametodo.

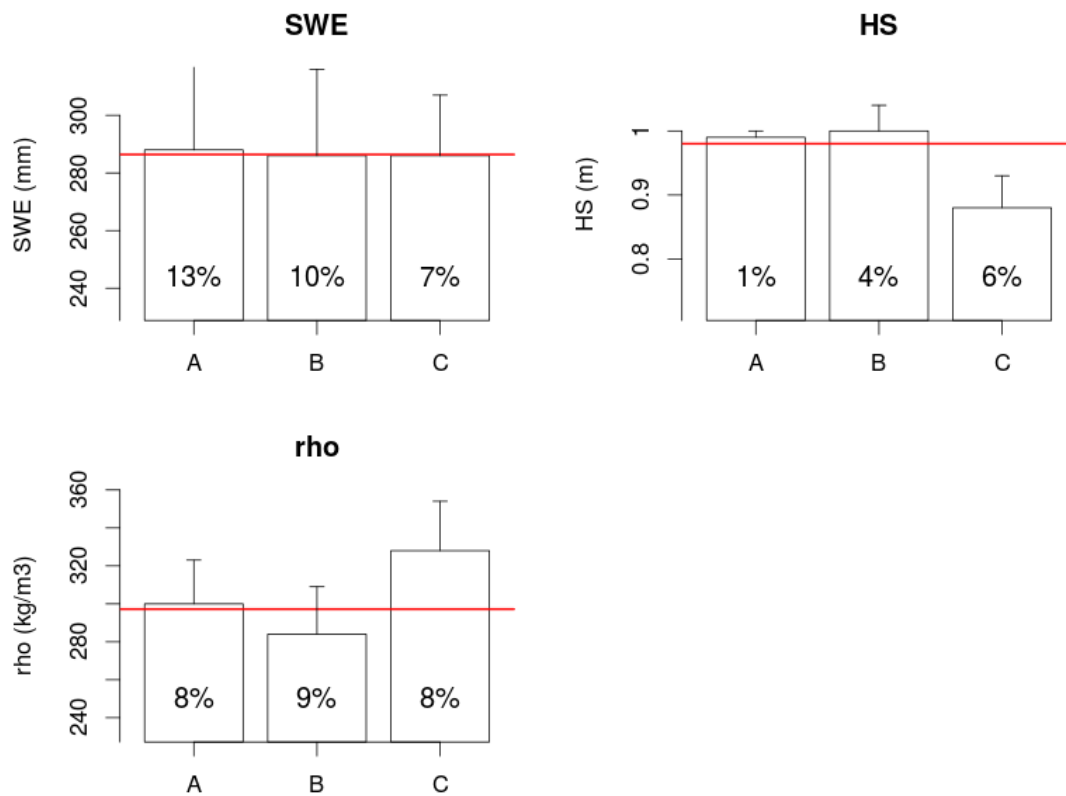


Figura 3: Valori mediani di SWE, HS e RHO in funzione del plot. La barra di errore rappresenta la mean absolute deviation (MAD). Si riporta inoltre il coefficiente di variazione, calcolato come $MAD/mediana * 100$. La linea rossa rappresenta la mediana generale.

Errore associato ai fattori esaminati

Il disegno sperimentale proposto permette di valutare simultaneamente la variabilità dell'SWE associata a 4 diversi fattori: i) all'errore della coppia di operatori (operatore), ii) alla variabilità tra coppie di operatori (interoperatore), iii) allo spazio e iv) al metodo utilizzato. La figura 4 mostra qual è l'errore medio associato a ciascuna di queste fonti di variabilità. L'errore medio associato all'operatore è del 2.5%. Ciò indica che a parità di condizioni del manto nevoso e di metodo utilizzato lo stesso operatore sbaglia, mediamente, del 2.5%. La distribuzione di frequenza dell'errore-operatore suggerisce inoltre che solo in pochissimi casi l'operatore commette un errore superiore al 5%. La variabilità interoperatore rappresenta la fonte di errore maggiore, intorno al 7%. Rappresenta la variabilità media tra osservazioni compiute da osservatori diversi nelle medesime condizioni e, pur mostrando valori relativamente bassi, è la fonte di variabilità maggiore dell'esperimento condotto. Si osservi a tal proposito la figura 5 che mostra tutti i dati di SWE misurati separati per plot e metodo. Analizzando la figura è evidente come la variabilità tra coppie determini variazioni nell'SWE anche dell'ordine di 100 mm. Bisogna tuttavia notare che all'interno della variabilità interoperatore potrebbe confluire anche la variabilità spaziale nello stesso plot, che assumiamo trascurabile lungo il transetto di circa 12-15 m che costituisce ciascun plot. La variabilità interoperatore potrebbe quindi essere sovrastimata.

La variabilità spaziale tra plot è minima (inferiore a 1%) come atteso data l'omogeneità spaziale dell'SWE nell'area osservata. La variabilità legata al metodo è la seconda per importanza e sarà oggetto di ulteriori considerazioni in seguito.

Fonte	Variabilità (%)
Operatore	2.5%
Interoperatore	7.1%
Spaziale (plot)	0.1%
Metodo	3.6%

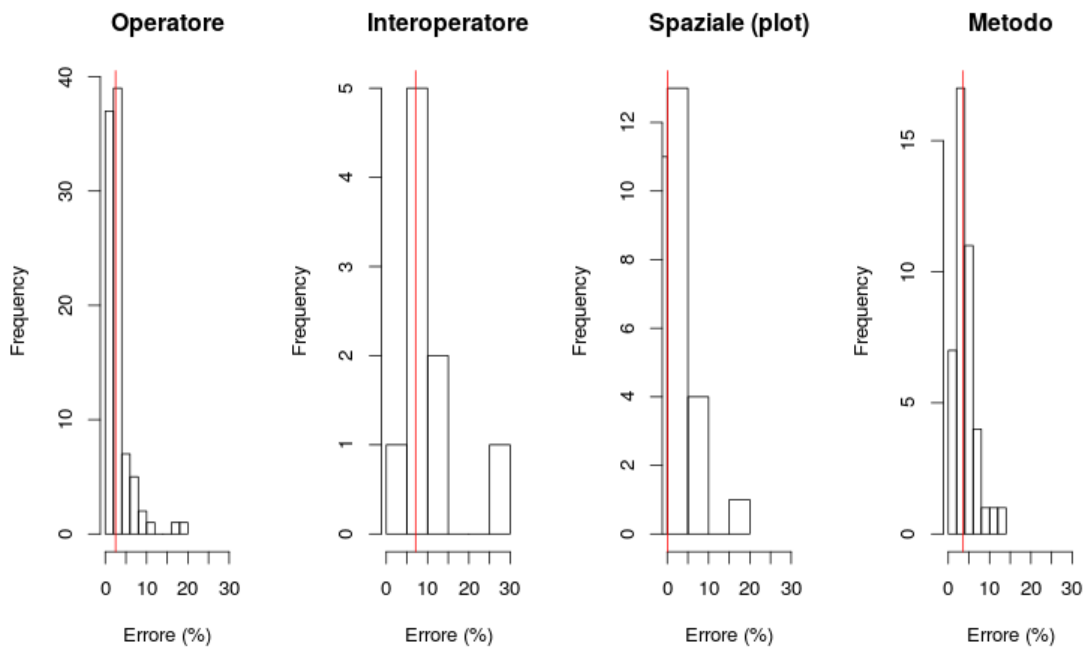


Figura 4: Errore percentuale medio associato a ciascuna delle tre fonti di variabilità e distribuzione di frequenza degli errori. La linea rossa rappresenta l'errore mediano, il cui valore è riportato in tabella

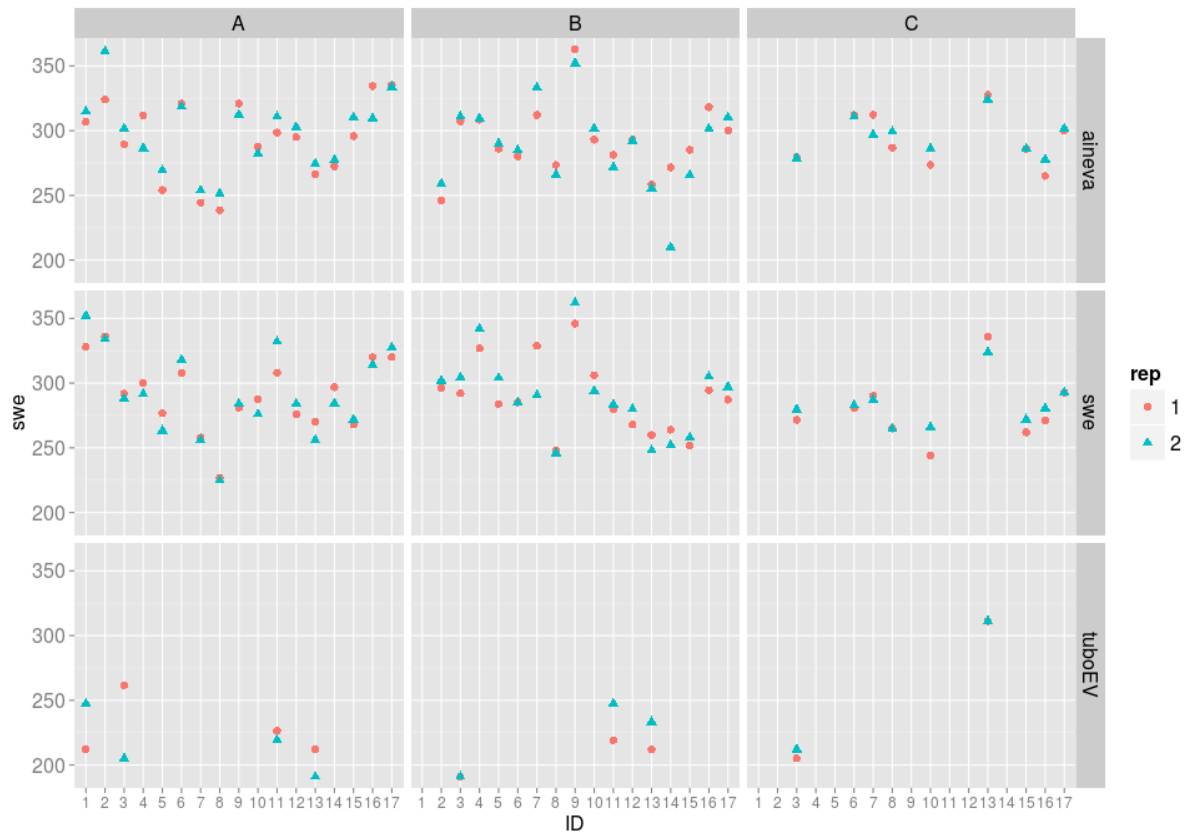


Figura 5: Tutte le misure di SWE effettuate, separate per plot e metodo. I colori diversi indicano le due repliche di ciascuna coppia

Variabilità tra metodi e spaziale

Per testare l'ipotesi che metodi diversi portino a differenze significative nella stima dell'SWE e che esista una variabilità spaziale significativa, è stato utilizzato un mixed model con metodo e plot (e la loro interazione) come fattori fissi e le coppie di misuratori come fattore

random. Questo modello permette di separare le fonti di variabilità legate al metodo, allo spazio, interoperatore ed intraoperatore al fine di testare qual è l'effetto delle singole fonti di variabilità sulle misure medie di SWE. I risultati in figura 6 dimostrano in maniera quantitativa ciò che si poteva ipotizzare già osservando la figura 2, cioè che la media delle osservazioni condotte col tubo EV è inferiore alle medie risultanti dagli altri due metodi. Le medie tra i diversi plot non sono invece statisticamente differenti.

Per analizzare ulteriormente i risultati del modello statistico, osserviamo (fig. 7) la tabella Anova del modello, che indica quali fattori hanno un effetto significativo su SWE. L'effetto metodo e la mancanza di effetto plot sono già stati discussi. E' interessante però osservare che esiste anche un effetto dovuto all'interazione tra metodo e plot. Ciò significa che l'effetto metodo non è lo stesso nei diversi plot. Per comprendere questo fenomeno si osservi il grafico riportato in figure 9 che mostra le medie (e variabilità) di SWE in funzione del metodo nei tre plot. La media del tubo EV è significativamente minore rispetto ai metodi SWE e Aineva nei plot A e B (così come rispecchiato dalle medie generali), mentre nel plot C lo scostamento tra metodi è inferiore e non significativo. Bisogna però considerare che la numerosità campionaria del plot C è decisamente inferiore rispetto ai primi due plot. Si tratta dunque di uno spunto interessante ma potenzialmente influenzato dalla diversa numerosità campionaria.

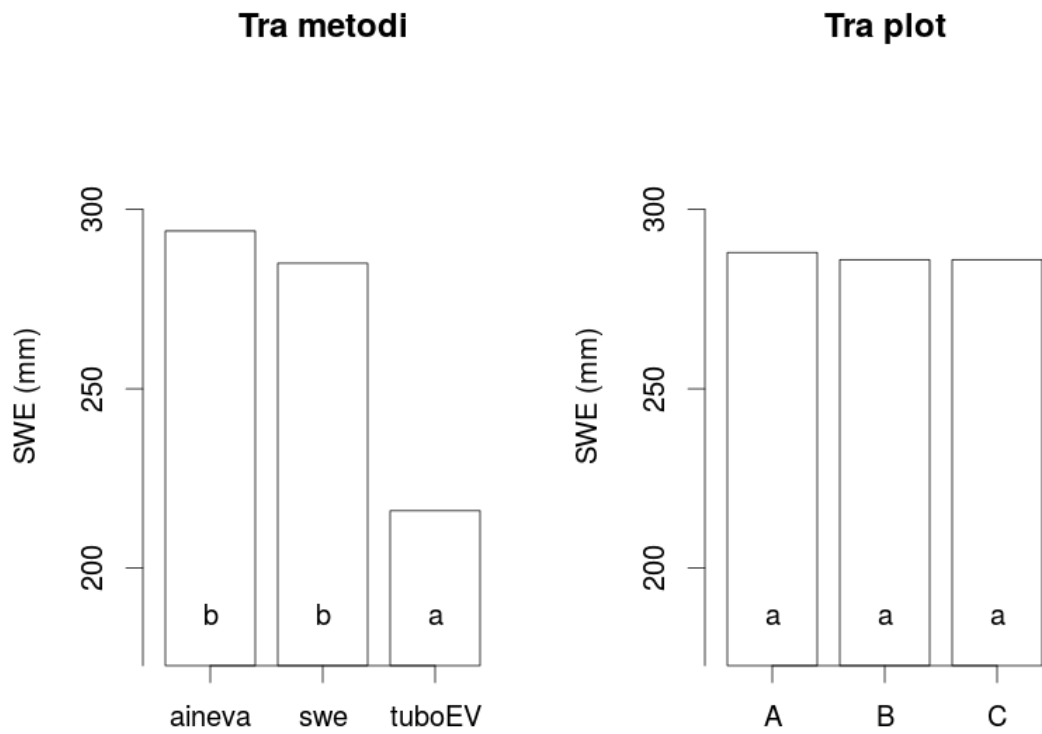


Figura 6: Risultati del modello statistico (mixed effect model). Le barre sono le stesse mostrate in fig. 2, mentre lettere diverse nella parte bassa della barra indicano differenze significative tra le medie

Figura 7: Tabella Anova del modello misto. p-values inferiori a 0.05 indicano un effetto significativo del fattore corrispondente. L'ultima riga (method:plot) si riferisce all'interazione tra i due fattori.

	F-value	p-value
method	37.089	0
plot	0.615	0.542
method:plot	2.7	0.033

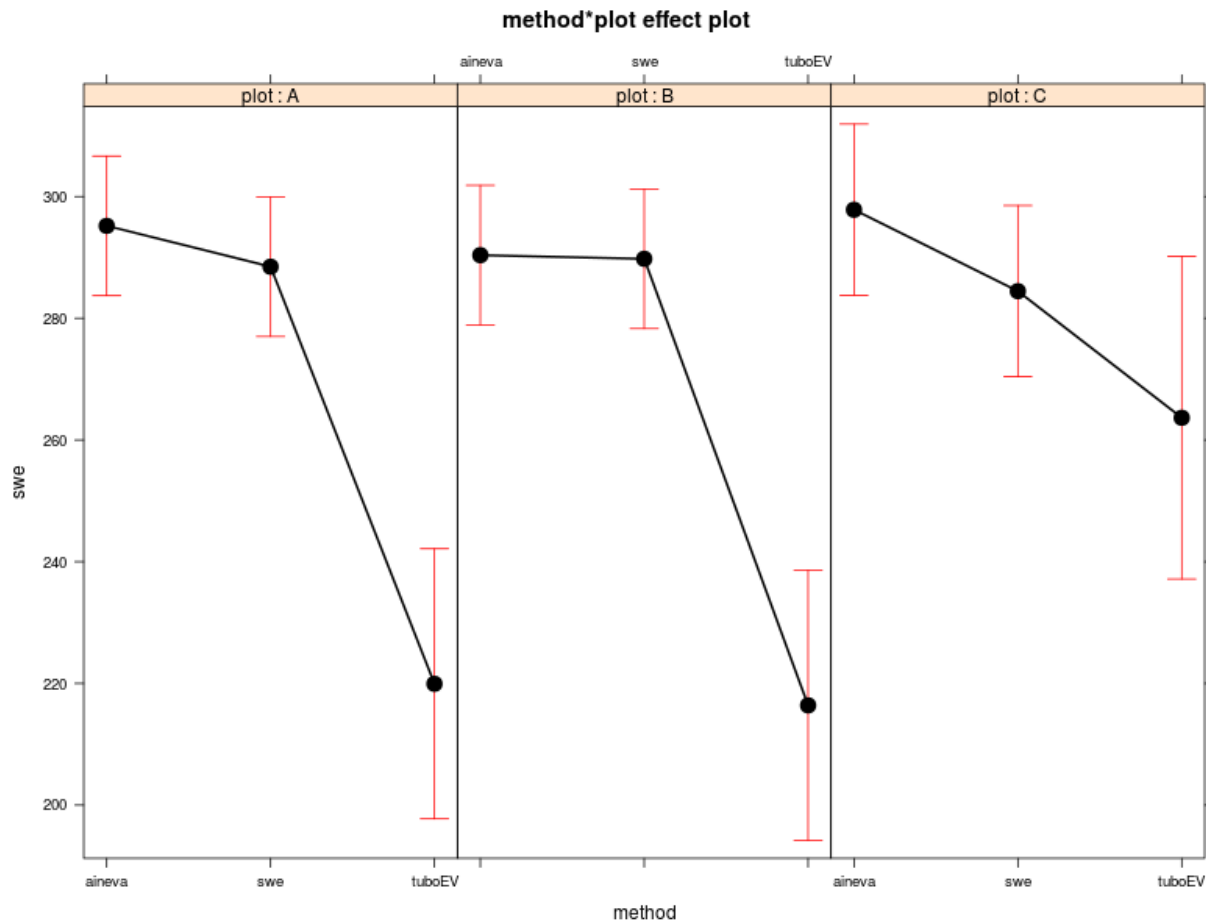


Figura 8: Medie di SWE separate per metodo e per plot

Appurato che i tre metodi portano a risultati diversi tra loro, ed in particolare che il metodo tuboEV offre una stima di SWE inferiore rispetto agli altri due metodi, è interessante testare se le medie ottenute con ciascun metodo differiscano da un valore di riferimento che possiamo considerare verità a terra. Non avanzando nessuna ipotesi a priori sulla maggiore

validità di un metodo rispetto ad un altro, la migliore stima della verità a terra che possiamo ottenere è la media di tutte le misure effettuate (SWE medio = 286 mm). La tabella seguente riporta il valore di significatività del test T necessario per verificare se i valori di SWE ottenuti coi tre metodi siano effettivamente diversi dalla verità a terra. Come atteso, le medie di SWE misurato coi metodi AINEVA ed SWE non differiscono dalla media generale, mentre la media del tubo EV è significativamente inferiore alla media generale (verità a terra).

method	pValue
aineva	0.083
swe	0.602
tuboEV	0.001

Figura 9: p-value del T test utilizzato per confrontare le medie di SWE dai singoli metodi con la verità a terra. p-values inferiori a 0.05 indicano una differenza significativa

Validità dei risultati ottenuti

Gli esiti dell'esperimento offrono interessanti spunti di discussione ma è giusto ricordarne i limiti. Al fine di testare la variabilità tra metodi si è cercato di trascurare la variabilità spaziale scegliendo un sito altamente omogeneo. I risultati ottenuti da questo esperimento vanno quindi estesi con cautela a condizioni molto diverse da quelle qui incontrate (topografia complessa, valori molto diversi di altezza e densità del manto nevoso).

L'elevata differenza nella numerosità campionaria delle misure SWE e AINEVA versus tuboEV potrebbero suggerire che le differenze riscontrate siano un artificio dovuto al basso numero di campioni ottenuti con il tubo EV. Per questo motivo abbiamo condotto un esperimento di ricampionamento casuale, in cui per 5000 volte abbiamo campionato un numero di misure di SWE e AINEVA pari al numero di campioni di tuboEV (cioè 18), al fine di simulare un uguale numero di campioni per ciascun metodo. Ne risulta che nel 99% dei casi le conclusioni sarebbero analoghe a quelle sopra esposte, cioè che il tubo EV porta ad una stima del SWE inferiore rispetto agli altri due metodi.

2 Conclusioni

In questa campagna di interconfronto sono state effettuate 186 misure su una superficie relativamente omogenea e subpianeggiante di circa 1000 m². Le misure sono state effettuate secondo tre metodi diversi ed in tre aree diverse. I metodi utilizzati sono: campionamento AINEVA con riconoscimento degli strati (denominato AINEVA), campionamento a profondità fisse (SWE), campionamento con carotatore (tubo EV).

Gli obiettivi dell'interconfronto erano quelli di:

- i) valutare la variabilità nelle misure di un singolo operatore in condizioni controllate.
- ii) valutare la variabilità tra operatori in condizioni controllate.
- iii) valutare la variabilità tra metodi.

I risultati mostrano che la variabilità del singolo operatore determina un errore nella misura inferiore al 3%, mentre la variabilità tra diversi operatori determina un errore intorno al 7%, rappresentando la fonte di variabilità maggiore. La variabilità spaziale dell'esperimento è risultata minima (inferiore al 1%).

I metodi AINEVA e SWE portano ad una stima dell'equivalente in acqua del manto nevoso confrontabile, mentre il metodo del tuboEV porta ad una sottostima dell'equivalente in acqua del manto nevoso.

Si conclude ricordando che l'elevata variabilità interoperatore osservata in questo esperimento in condizioni altamente omogenee pone l'accento sull'utilità e sull'ulteriore necessità di esercizi di interconfronto di questo tipo, al fine di uniformare il più possibile la tecnica di campionamento e la realizzazione delle misure.

3 Allegati

Per avere una panoramica di tutte le misure e non solo dei dati aggregati, presentiamo due grafici analoghi a quello di figura 5, ma riferiti alle misure di RHO ed HS effettuate da ciascuna coppia.

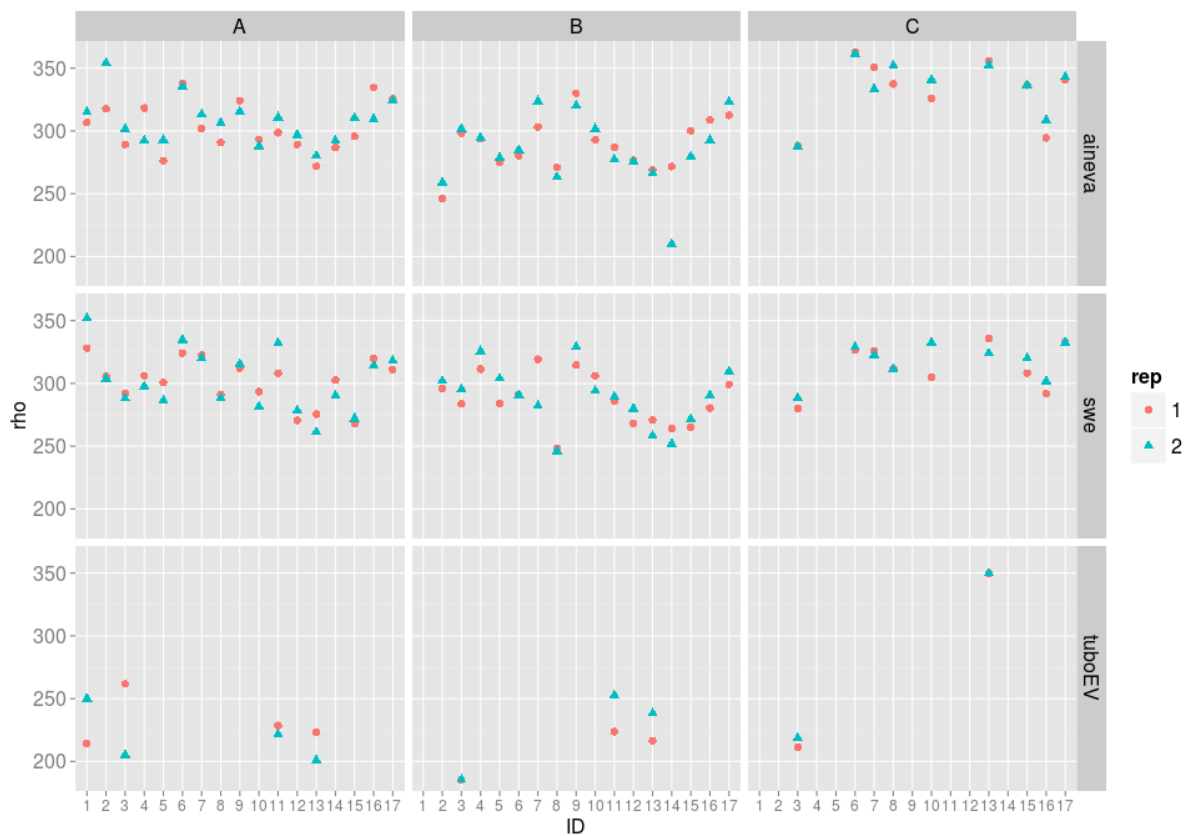


Figura 10: Tutte le misure di densità del manto nevoso (media ponderata), separate per plot e metodo. I colori diversi indicano le due repliche di ciascuna coppia

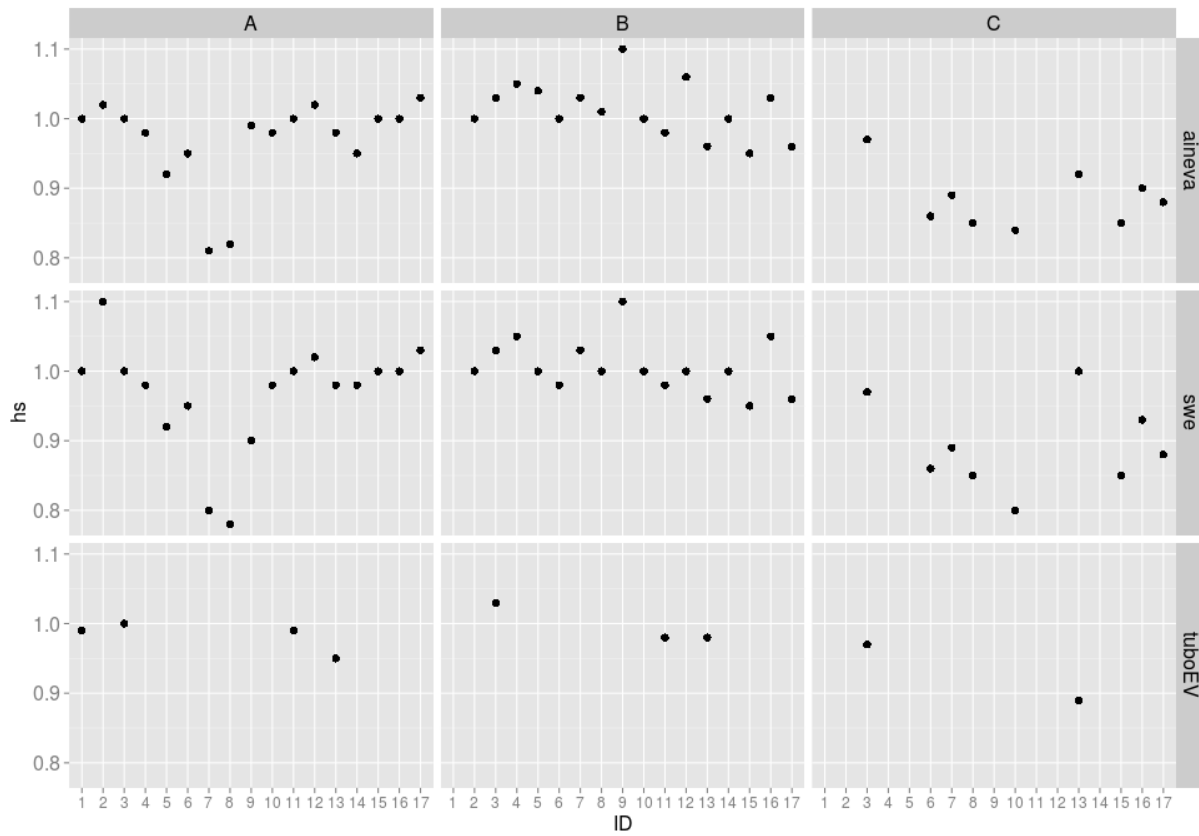


Figura 11: Tutte le misure di HS effettuate, separate per plot e metodo.