

STIMA della NEVE FRESCA dai DATI delle STAZIONI AUTOMATICHE

Definizione di un modello per migliorare la valutazione dell'altezza di neve fresca dalla lettura dei dati delle stazioni automatiche

**Giulia Rossi e
Mattia Faletto**
con la collaborazione di
**Maria Cristina Prola e
Roberto Cremonini**

ARPA Piemonte, Dipartimento
Sistemi Previsionali, Torino
mariacristina.prola@arpa.piemonte.it

Il quantitativo di neve fresca, misura in centimetri della neve caduta nelle ultime 24 ore, HN, è di fondamentale importanza per la valutazione del pericolo valanghe e delle condizioni di stabilità del manto nevoso. Il dato di neve fresca derivato dalle frequenti misure di altezza del manto nevoso effettuate dalle stazioni nivometriche automatiche risulta tendenzialmente sottostimato. Questa sottostima è causata *in primis* dai processi di assestamento del manto nevoso preesistente alla nevicata a causa del peso della nuova neve. Le osservazioni nivologiche manuali effettuate quotidianamente dagli operatori (Modello 1 - AINEVA) sono affette da questa sottostima in modo molto limitato, poiché la misura viene eseguita tramite l'utilizzo di una tavoletta di legno, posta sulla superficie del manto nevoso, che ha funzione di separare il processo di assestamento del manto nevoso dall'accumulo della nevicata.

Lo scopo di questo studio è stato quindi quello di creare un algoritmo di correzione dei valori di neve fresca restituiti dalle stazioni automatiche, utilizzando come riferimento i valori di neve fresca modellizzati da SNOWPACK (Bartelt and Lehning, 2002), modello numerico unidimensionale sviluppato dall'Istituto Federale Svizzero (SLF) che simula l'evoluzione temporale del manto nevoso, tenendo conto dei molti fattori nivometeorologici che lo influenzano



REGIONE PIEMONTE
AZIONE METEOROLOGICA
del CHIOT

INTRODUZIONE

Una delle principali funzioni delle stazioni nivometriche automatiche risiede nell'elevata frequenza temporale di raccolta dei dati relativi all'altezza della neve al suolo, HS, (mediamente ogni 30'), che permette di monitorare quasi in tempo reale l'evoluzione di una precipitazione nevosa. La misura automatica della neve avviene attraverso un sensore ad ultrasuoni, costituito da una coppia di trasduttori ultrasonici in aria che misurano il tempo di percorrenza dell'impulso dalla trasmissione alla ricezione, dopo essere stato riflesso dalla superficie del manto nevoso. Tale valore viene compensato con il dato di temperatura dell'aria, con margine di incertezza di 1,5 cm, tollerato fino a 4 cm. Questo significa che è possibile avere un *range* di misura di 5,5 cm. Nonostante la compensazione, il grafico delle HS riporta oscillazioni anche di 10 cm, in particolare in corrispondenza di brusche variazioni di temperatura che si possono registrare nell'arco diurno nelle giornate di tempo sereno e stabile, ma anche durante i passaggi di fronti con precipitazioni associate (Frigo et al., 2012). Tale misura può risultare alterata anche da fattori, occasionali, quali interferenze dovute ad ostacoli temporanei che interferiscono con la misura del sensore. (Terzago et al., 2012).

La Regione Piemonte è dotata di una rete di stazioni nivometriche automatiche per il monitoraggio degli apporti nevosi e di conseguenza per la valutazione del pericolo valanghe e di altri aspetti in campo idrologico. Data l'importanza di queste attività, è necessario che i dati delle stazioni automatiche siano affidabili e accurati. Ciò non è però sempre possibile: le misurazioni automatiche sono spesso soggette ad errori per varie cause e fattori, legati perlopiù alle condizioni meteorologiche; per esempio, si possono registrare accumuli o riduzioni repentine di neve a causa di forti venti oppure si può verificare una compattazione della neve più o meno accentuata in funzione delle condizioni climatiche (McClung and Schaerer, 1993), la quale influisce sulle misurazioni, renden-

do sottostimata l'altezza di neve fresca. Per questi motivi, si è sviluppato il "Data Quality Control" (DQC), ossia un insieme di sistemi di controllo qualità dei dati ricevuti; essi consistono in tutte quelle procedure operative che saranno effettuate durante il normale funzionamento del sistema di monitoraggio per garantire un corretto funzionamento (Salvati and Brambilla, 2008).

Arpa Piemonte, oltre a tarare periodicamente gli strumenti, affianca alle misurazioni automatiche quelle manuali, effettuate da personale appositamente formato, che forniscono dati nivometeorologici orientati alla "qualità" e distribuzione della neve. I dati delle stazioni automatiche insieme a quelli manuali vengono analizzati da un operatore che effettua controlli giornalieri. Il processo di verifica necessita di migliorie che riducano la componente soggettiva del controllo, adottando processi automatizzati su base di algoritmi matematici. In questo studio si è cercato di creare un modello di correzione dei valori di neve fresca restituiti dalle stazioni automatiche in modo da ridurne la sottostima. Per fare questo si sono utilizzati i dati di neve fresca valorizzati dal modello SNOWPACK (Bartelt and Lehning, 2002).

Il lavoro si basa sul confronto tra i valori restituiti dalle stazioni automatiche e quelli restituiti da SNOWPACK per l'inverno 2013-2014, considerando gli ultimi come veritieri e correggendo i primi in funzione di essi. Utilizzando una regressione lineare si è quindi definito un algoritmo che, utilizzando alcuni parametri noti, corregge la stima di neve fresca ottenuta dalle stazioni automatiche. L'accuratezza del modello è stata poi valutata applicandolo ai valori dell'inverno precedente (2012-2013) e confrontando i valori "corretti" con quelli definiti in SNOWPACK.

MATERIALI E METODI

Arpa Piemonte dispone su suo territorio di competenza di più di 100 stazioni nivometriche automatiche; tra queste ne sono scelte 5, a quote che vanno da 1875 m s.l.m. a 2820 m s.l.m., di cui sono di-

sponibili anche i risultati di SNOWPACK. Le stazioni prese in considerazione sono Limone Piemonte (1875 m s.l.m.), Passo del Moro (2820 m s.l.m.), Bocchetta della Pisse (2410 m s.l.m.), Sestriere-Banchetta (2480 m s.l.m.) e Lago Pilone (2280 m s.l.m.). Per ognuna di esse si sono considerati i valori di HN (altezza della neve fresca) stimati dalle stazioni automatiche e da SNOWPACK, graficati poi insieme per evidenziarne le differenze. Ogni punto rappresenta una coppia di misurazioni: maggiore è la sua distanza dalla bisettrice, maggiore è la differenza tra i due valori considerati (Figg. 1a, 1b, 1c, 1d, 1e). In particolare, siccome in ascissa si hanno le altezze misurate dalle stazioni automatiche e sulle ordinate i valori corrispondenti in SNOWPACK, se i punti sono spostati verso l'asse orizzontale significa che il valore restituito dalla stazione automatica è sovrastimato (Fig. 1d), viceversa se i punti sono spostati più verso l'asse verticale allora le stazioni automatiche hanno sottostimato la grandezza (Fig. 1b, Fig. 1e).

Dai grafici si può notare come in alcuni casi ci siano differenze tra il valore della neve fresca della stazione automatica e quello di SNOWPACK che raggiungono anche 60 cm (Fig. 1a); pertanto si è voluta comprendere la causa di tali diversità. Controllando i dati archiviati nei database di Arpa Piemonte si è potuto vedere che le suddette differenze erano dovute al fatto che l'altezza di neve fresca registrata risultava alterata per motivi di attività eolica forte, la quale creava accumulo o diminuiva lo spessore reale di neve fresca precipitata. Tenendo invece in considerazione questo fenomeno, SNOWPACK restituiva valori elevati più simili a quelli misurati; in questi casi si è deciso quindi di riportare i valori delle stazioni automatiche alla stima iniziale che considerava l'azione eolica, in quanto si è trattato non di errori di misura, ma di un diverso approccio tra SNOWPACK e l'archiviazione dei dati in Yeti.

Apportando quindi queste correzioni, si è studiata la serie di dati relativa alla stagione invernale 2013-2014 tramite il

supporto del software *R Cran* (R, 2010). Innanzitutto, si sono presi in considerazione alcuni fattori che si ipotizzava inizialmente influissero sui valori registrati di neve fresca: la temperatura massima,

minima e media dell'aria, la temperatura superficiale della neve, l'altezza di neve fresca cumulata in uno, due, tre e quattro giorni. Successivamente si è valutata l'importanza relativa di questi parametri:

per farlo è stata definita la correlazione tra essi e l'altezza di neve fresca. Come si può notare in Fig. 2, si è ottenuta una correlazione significativa con i seguenti parametri: AirT (temperatura media dell'aria),

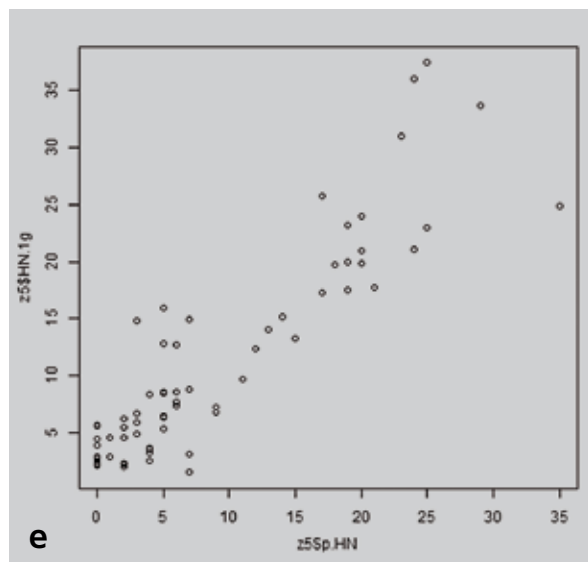
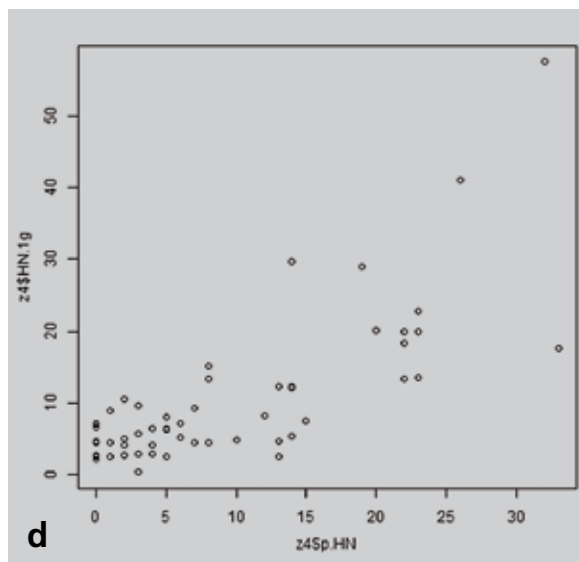
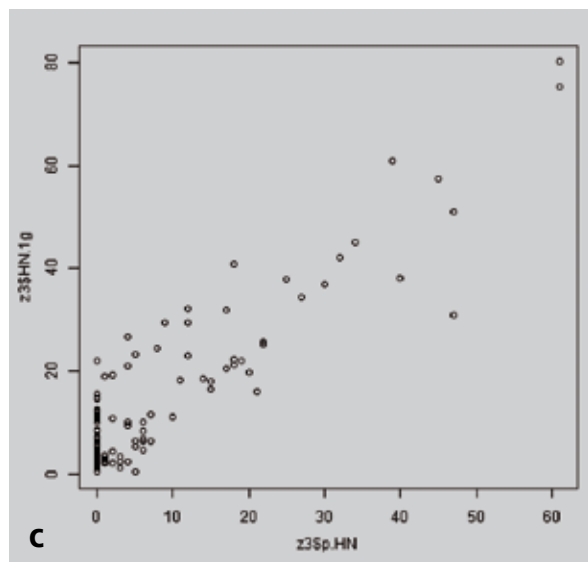
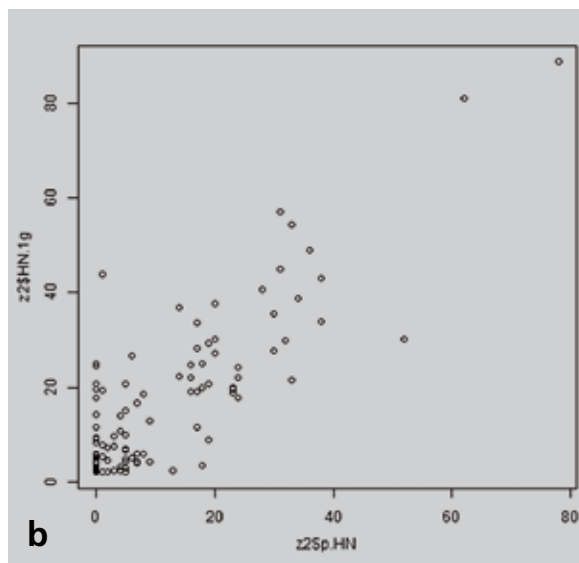
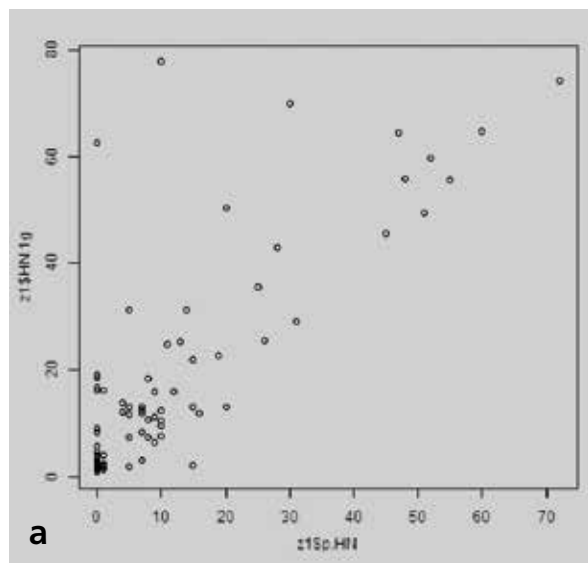


Fig. 1 - Valori di HN misurati dalle stazioni automatiche (in ascissa) e stimati da SNOWPACK (in ordinata):
a) Limone Piemonte,
b) Passo del Moro,
c) Bocchetta della Pisse,
d) Lago Pilone.



Tmax (temperatura massima giornaliera), HN, HN2g, HN3g, HN4g (altezza di neve cumulata in 1, 2, 3 e 4 giorni). Considerando queste variabili, si sono voluti effettuare ulteriori test statistici per individuare con maggiore precisione l'importanza relativa dei singoli fattori: col pacchetto "relaimpo" (U. Groemping and M. Lehrkamp, 2013), si sono effettuati calcoli delle relative importanze secondo 4 metodi:

1- **lmg**: rappresenta il contributo pesato dei regressori a R^2

2- **last**: rappresenta il contributo di ogni variabile se considerata per ultima
 3- **first**: è il contributo di ogni variabile se considerata per prima

4- **pratt**: indica il contributo del regressore proporzionalmente alla relativa correlazione col parametro di riferimento.

In Fig. 8 si possono osservare i risultati ottenuti con ciascun test: in R si sono effettuati i 4 test statistici. Il risultato è simile sia che si analizzino le singole stazioni separatamente che nell'insieme. A titolo

di esempio si riportano i grafici dei test che raggruppano tutte le stazioni (Fig. 3). Come si può notare i test indicano che le variabili più significative sono HN, HN2g e HN3g, ossia l'altezza di neve di uno, due e tre giorni; fa eccezione l'ultimo test che riporta risultati differenti, ma essendo in letteratura definito come il meno affidabile si è deciso di trascurarlo.

Si è scelto quindi di procedere con regressione lineare semplice e poi multipla per definire un modello col quale correggere le altezze di neve misurate dalle stazioni automatiche.

IL MODELLO

Lo scopo del modello definito è la correzione dei valori elaborati dalle stazioni automatiche in modo da avvicinarli ai valori reali, considerando rappresentativi quelli di SNOWPACK, con un margine di errore accettabile. Per fare ciò si è iniziato dal modello più semplice: la regressione lineare semplice, stimando il valore atteso dell'altezza di neve fresca dati i valori misurati dalla stazione automatica. L'equazione rappresentativa è dunque la seguente:

$$HN_{reale} = \alpha + \beta \cdot HN_1$$

Dove HN_1 è il valore misurato dalla stazione automatica relativo alla neve fresca di un giorno, ed i coefficienti α e β sono gli

Fig. 2 - Correlazioni dell'altezza di neve fresca con altri parametri.

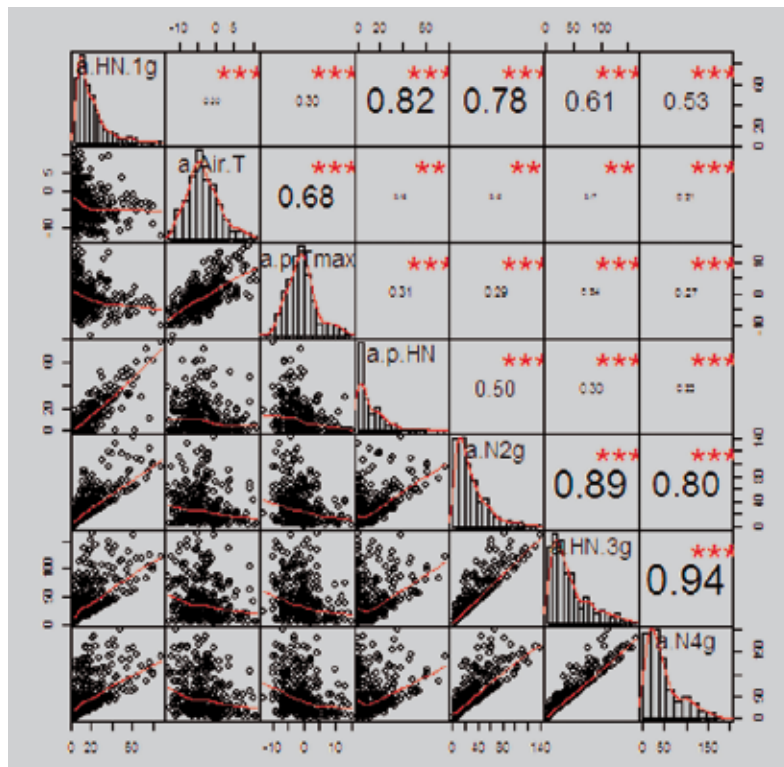
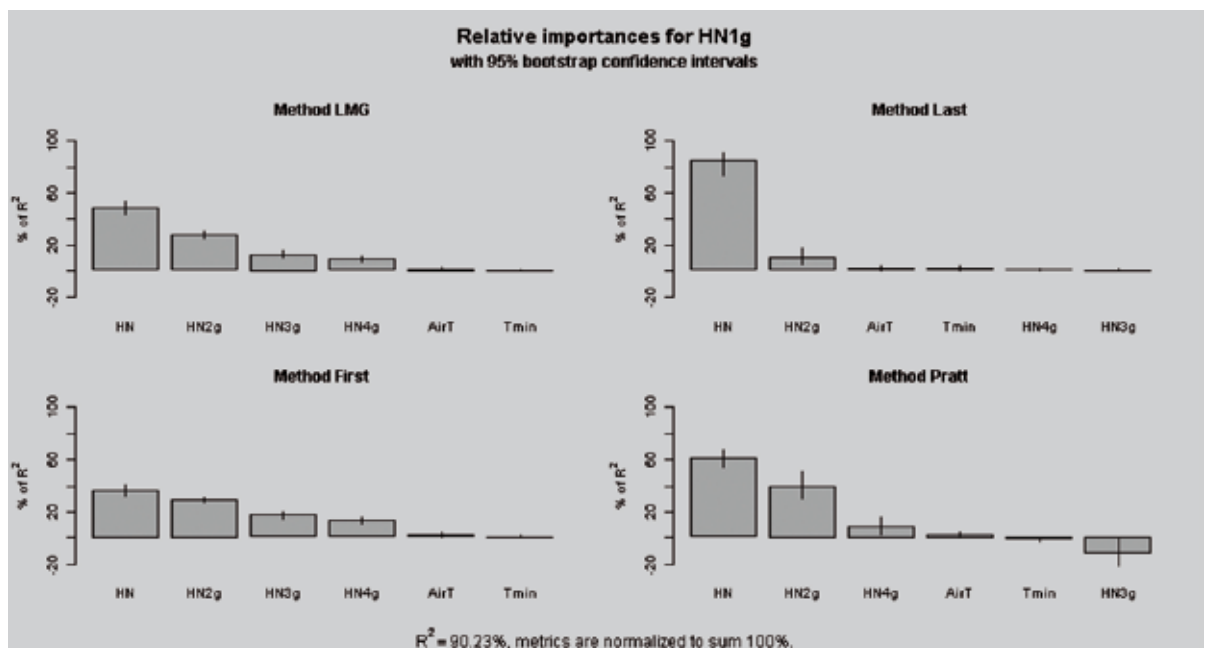


Fig. 3 - Importanza relativa dei parametri considerati.





stimatori del modello da ricercare. Svolgendo la regressione in R, considerando in un'unica volta tutti i valori di tutte le stazioni per avere una base più ampia, si sono ottenuti i seguenti risultati:
 $\alpha = 6,11$
 $\beta = 0,938$
 $R^2 = 0,76$

Il risultato non è molto soddisfacente, in quanto il valore di R^2 è piuttosto basso ed i residui massimo e minimo sono 36 ([cm]) e -25 ([cm]), cioè abbastanza elevati; si è dunque deciso di procedere con la regressione lineare multipla, ossia di aggiungere le altre variabili oltre ad HN1. In questo caso si ha:

$$HN_{reale} = \alpha + \beta \cdot HN1 + \gamma \cdot HN2 + \delta \cdot HN3$$

Questa volta gli stimatori da definire sono 4, e si hanno due variabili in più: HN2 (altezza di neve fresca cumulata in 2 giorni) e HN3 (altezza di neve fresca cumulata in 3 giorni). In questo modo il modello dovrebbe essere più preciso, in quanto il valore finale della variabile HNreale è stimato tenendo conto dell'influenza di 3 parametri anziché di uno solo. I risultati

sono i seguenti:

$$\alpha = 1,25$$

$$\beta = 0,685$$

$$\gamma = 0,251$$

$$\delta = 0,007$$

$$R^2 = 0,9$$

Si può notare come il valore di R^2 sia ora di 0,9 ed i residui massimo e minimo siano scesi rispettivamente a 14 ([cm]) e -21 ([cm]); dunque l'affidabilità del modello è accresciuta notevolmente. Pertanto tra la regressione lineare semplice e quella multipla, è più accurata e adatta la seconda. Per completezza si è voluto poi valutare un modello di regressione con legame quadratico tra la variabile HNreale e le altre, per comprendere se in tal modo i dati erano meglio rappresentati. In questo caso la relazione, considerando una sola variabile, è del tipo:

$$HN_{reale} = \alpha + \beta \cdot HN1 + \gamma \cdot (HN1)^2$$

I risultati, come nel caso dove si considerano tutte e tre le variabili (HN1, HN2, HN3), indicano valori residui e di R^2 praticamente identici alle regressioni precedentemente illustrate. Ciò significa che i

termini quadratici non forniscono apporti positivi al risultato finale; pertanto, a parità di "prestazione" si sceglie il modello con meno complessità (ossia quello senza legami quadratici) e quindi con ridotti oneri di calcolo. Non si è proceduto con regressioni di grado superiore, in quanto quasi sempre l'interpretazione dell'equazione di regressione è tanto più attendibile e generale quanto più la curva è semplice, come quelle di primo o di secondo grado. Regressioni di ordine superiore sono quasi sempre legate alle variazioni casuali; sono effetti delle situazioni specifiche del campione raccolto e solo molto raramente esprimono relazioni reali, non accidentali, tra le variabili (Soliani, 2005).

VALUTAZIONE DEL MODELLO

Il modello è stato successivamente verificato: per fare ciò si sono applicati i coefficienti di regressione ai dati relativi all'inverno dell'anno 2012-2013, precedente a quello su cui si è costruito il modello. Si è applicato il modello ai dati delle stazioni automatiche dell'inverno precedente, escludendo chiaramente i giorni in

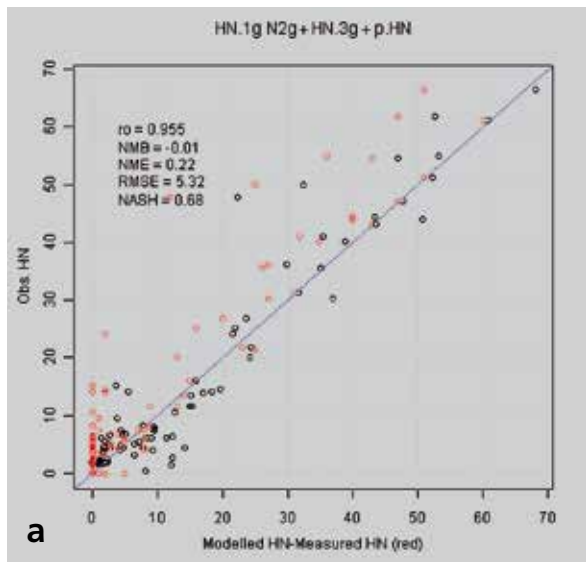
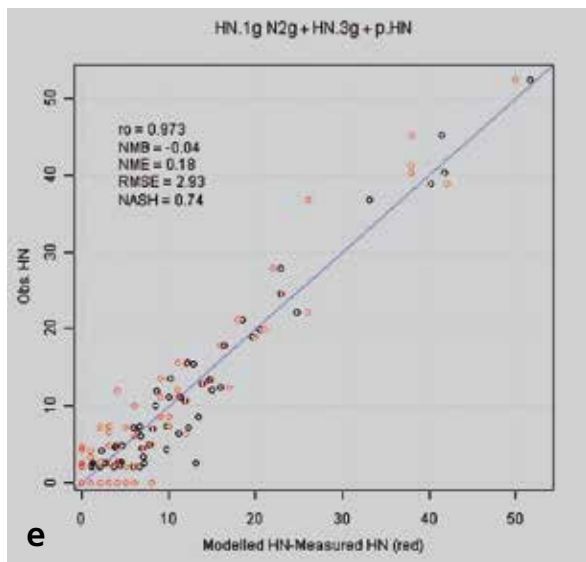
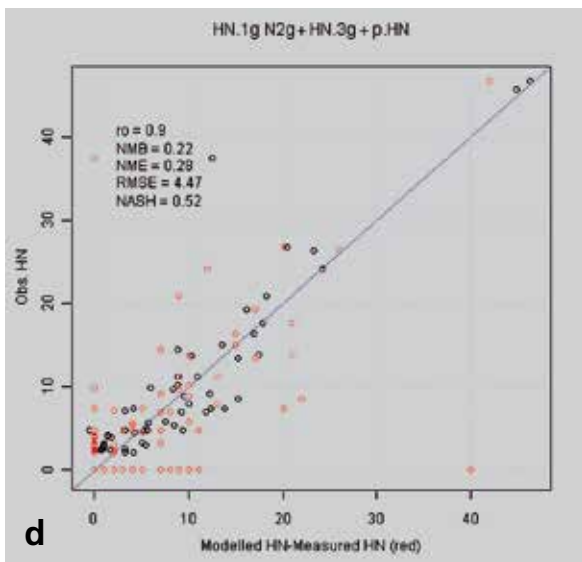
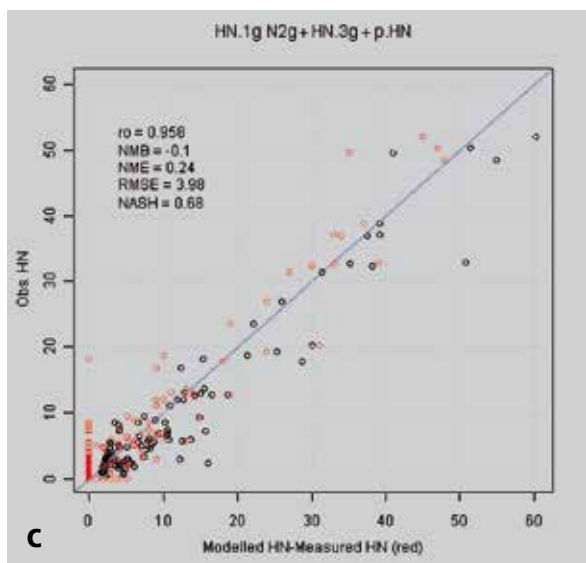
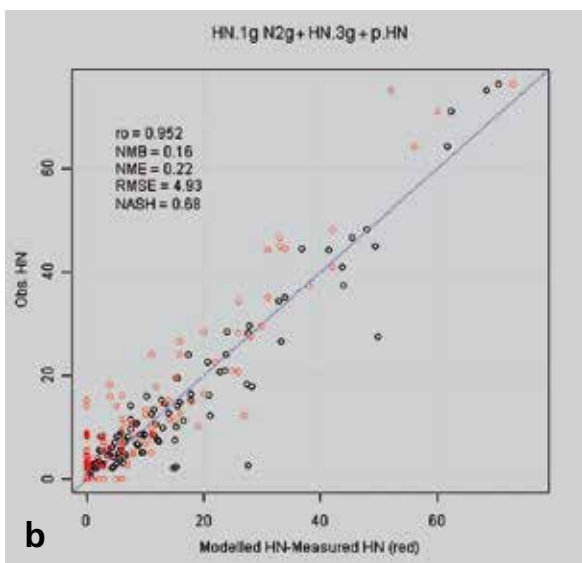


Fig. 4- Modello applicato alle diverse stazioni:
 a) Limone Piemonte,
 b) Passo del Moro,
 c) Bocchetta della Pisse,
 d) Lago Pilone.



cui non è nevicato: al valore dell'altezza di neve fresca cumulata di un giorno si è applicata la "correzione", legata ai parametri di altezza di neve fresca cumulata di 1, 2 e 3 giorni misurata dalle stazioni

automatiche. Ottenuti i dati corretti, essi sono stati confrontati con i valori di SNOWPACK, per valutare la qualità della correzione apportata e quale margine di errore è ancora possibile. Il procedimento

è stato effettuato sia sulle singole stazioni che sull'insieme di tutti i dati di tutte le stazioni unitamente.

Nei grafici riportati sono rappresentati in rosso i dati non corretti (cioè sem-

plicemente i dati restituiti dalle stazioni automatiche), in nero quelli corretti, ossia quelli ottenuti dopo l'applicazione del modello, la linea blu indica invece la regressione. In Fig. 4 (a, b, c, d, e) sono rappresentati i risultati per le singole stazioni, rispettivamente Limone Piemonte, Passo del Moro, Bocchetta della Pisse, Banchetta e Lago Pilone; in Fig. 5 invece sono rappresentati in un solo grafico i dati di tutte le stazioni.

Come si può notare i dati corretti in tutti i casi si spostano chiaramente verso la linea di regressione, ossia le correzioni tendono effettivamente a far coincidere il dato registrato dalle stazioni automatiche con quello derivato da SNOWPACK. Per quanto riguarda la precisione e l'affidabilità del modello, esso presenta nelle varie stazioni un errore residuo (NME) compreso tra 18% e 28% per una media di circa il 22%, come ottenuto nel caso in cui sono stati uniti tutti i dati; infine si ha un valore di R^2 pari a 0,95, che quindi indica una proporzione tra la variabilità dei dati e la correttezza del modello statistico utilizzato abbastanza buona.

Il modello è stato poi applicato agli stessi dati, ma considerando solo gli eventi caratterizzati da una precipitazione nevosa superiore ai 3 cm: infatti studi passati hanno messo in luce che la variabilità delle misure di spessore del manto nevoso è caratterizzata da un range di 3 cm, imputabile alla precisione dello strumento (Frigo et al., 2012). Per questo motivo, escludendo eventi con precipitazioni minori di 3 cm, si è cercato di eliminare l'errore strumentale; i risultati sono comunque molto simili, con un R^2 pari a 0,96 ed un errore residuo del 20%.

CONCLUSIONI

Lo scopo del lavoro svolto è stato quello di fornire un supporto alla validazione dei dati provenienti dalle stazioni automatiche, tramite la creazione di un modello che corregga l'altezza di neve fresca registrata, in funzione della neve cumulata in uno, due e tre giorni consecutivi.

Per costruire il modello si è analizzato l'inverno 2013-2014. Sono stati presi i dati

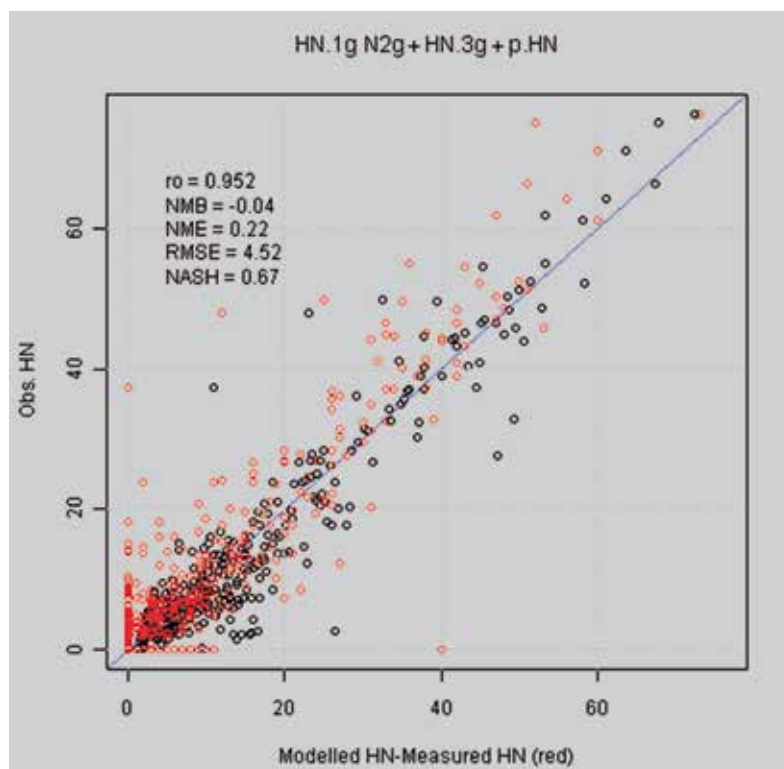


Fig. 5 - Modello applicato all'insieme delle stazioni.

delle stazioni automatiche e confrontati con quelli del programma SNOWPACK, considerandoli reali; si è elaborata una regressione lineare che rappresentasse i dati disponibili. Il modello che è stato definito, presenta un R^2 pari a 0,9 e residui massimi e minimi di 14 e -21; esso è stato quindi testato sull'inverno precedente, 2012-2013, applicandolo ai dati delle stazioni automatiche e valutando poi i valori ottenuti con quelli reali. Ciò che si è ottenuto è che il modello, considerando una media su tutte le stazioni, presenta un errore residuo pari a circa il 22% (o 20% non considerando gli eventi di precipitazione minori di 3 cm): è pertanto un buon risultato che migliora di molto i valori restituiti dalle stazioni automatiche, che senza l'applicazione del modello presentano valori che si discostano da quelli reali anche dell'80-90%.

La correzione apportata dal modello fornisce sicuramente un miglioramento nella stima del valore di neve fresca misurato dalle stazioni automatiche. Grazie all'utilizzo delle informazioni di SNOWPACK (molto complesse e dettagliate) si è potuto ottenere uno strumento semplice con un miglioramento della qualità dei dati.

Il modello elaborato è sicuramente un buon inizio. E' auspicabile un ulteriore

affinamento del modello, magari individuando ulteriori parametri che permettano di perfezionare la correzione.



Bibliografia

- Bartelt P., Lehning M. (2002) - A physical SNOWPACK model for the Swiss avalanche warning. WSL, Swiss Federal Institute for Snow and Avalanche Research, SLF, Davos, Svizzera.
- Frigo B., Prola M.C., Faletto M. (2012). Valutazione della stabilità del manto nevoso: linee guida per la raccolta e l'interpretazione dei dati, Regione Autonoma Valle d'Aosta.
- Goemping U., Lehrkamp M. (2013) - Relative importance of regressors in linear models. R Foundation for Statistical Computing, Vienna, Austria.
- McClung D., Schaerer P. (1993) - Manuale delle Valanghe. Zanichelli
- R Development Core Team. (2010) - R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Salvati M., Brambilla E. (2008) - Data quality control procedures in Alpine meteorological services. FORALPS Technical Report. Università degli studi di Trento, Dipartimento di Ingegneria Civile e Ambientale, Trento
- Soliani L. (2001) - Fondamenti di statistica applicata all'analisi e alla gestione dell'ambiente. <http://www.chierici.biz/biblio/statistica%20ambientale.pdf>
- Terzago S., Prola M.C., Frantianni S., Cremonini R., Barbero S. (2012) - Validazione di misure nivometriche automatiche. Neve e Valanghe, n° 75 Aprile 2012. Aineva.