Ševčík, J., 2026. Snow data assimilation in the ALADIN model. In: Tolasz, R., Šustková, V. (Ed.), Proceedings of the 2nd PERUN Project Conference (TA ČR, SS02030040). CHMI, Praha, s. 69–74, ISBN 978-80-7653-095-9. https://doi.org/10.59984/978-80-7653-095-9.09

# Snow data assimilation in the ALADIN model

## (Asimilace měření sněhu v modelu ALADIN)

Ševčík, J., Trojáková A., Meier, F., Brožková, R., Žák, M.

https://doi.org/10.59984/978-80-7653-095-9.09

Jáchym Ševčík, Czech Hydrometeorological Institute; Faculty of Mathematics and Physics, Charles University, jachym.sevcik@chmi.cz

Alena Trojáková, Czech Hydrometeorological Institute, alena.trojakova@chmi.cz

Florian Meier, GeosSphere Austria, Department Analyse und Modellentwicklung, florian.meier@geosphere.at

Radmila Brožková Czech Hydrometeorological Institute, radmila.brozkova@chmi.cz

Michal Žák, Faculty of Mathematics and Physics, Charles University, michal.zak@matfyz.cuni.cz

***Abstract:*** Snowpack plays an important role in the climate system, and it is thus desirable to represent it correctly in any numerical weather prediction or climate model. Such representation can be effectively improved by making use of available observations via the process of data assimilation. This contribution describes the deployment of a snow depth assimilation framework in the ALADIN system at the CHMI (Czech Hydrometeorological Institute). The assimilation algorithm used is the statistical interpolation method. Three distinct variants of the scheme are proposed to test various approaches towards error covariance estimation and high-altitude quality control. The parameters of these variants are then optimized using the Nelder-Mead simplex minimization algorithm. Such optimized settings result in significantly improved snow depth forecast performance compared to a reference without the assimilation scheme. The impact on other surface variables such as screen-level temperature or relative humidity is mixed, and a means of mitigating the negative effects is described and tested. A more detailed description of the implementation and results appears in an article submitted to the Meteorological bulletin.

***Keywords:*** snow cover – statistical interpolation – snow fraction

***Abstrakt:*** Sněhová pokrývka hraje v klimatickém systému významnou roli a je tak žádoucí ji správně reprezentovat v jakémkoli numerickém předpovědním nebo klimatickém modelu. Tuto reprezentaci lze efektivně zlepšit využitím dostupných pozorování prostřednictvím asimilace dat. Tento příspěvek popisuje zprovoznění nového systému asimilace výšky sněhu v modelu ALADIN v ČHMÚ (Českém hydrometeorologickém ústavu). Použitým asimilačním algoritmem je metoda optimální interpolace. Jsou testovány tři odlišné varianty schématu, za účelem otestování různých přístupů k odhadu kovariancí chyb a kontrole kvality ve vysokých nadmořských výškách. Parametry všech konfigurací jsou následně optimalizovány pomocí simplexového minimalizačního algoritmu. Takto optimalizovaná nastavení vedou k výrazně lepší kvalitě předpovědí výšky sněhu ve srovnání s referencí bez asimilačního schématu. Dopad na další přízemní proměnné, jako je teplota ve 2 metrech nebo relativní vlhkost, je smíšený, přičemž je popsán a otestován způsob zmírnění negativních dopadů. Detailnější popis implementace a výsledků pak lze najít v článku předloženém k recenznímu řízení časopisu Meteorologické zprávy.

***Klíčová slova:*** sněhová pokrývka – optimální interpolace – sněhová frakce

## Introduction

When present, snowpack is a critical element of the climate system with very specific properties, such as high albedo and thermal emittance, the ability to insulate the ground from the atmosphere and to store water. Through these properties, the presence of snowpack influences the surface radiative, energy and hydrological balance. (Barry 2002). For this reason, snow needs to be accounted for when preparing the initial conditions for both numerical weather prediction forecasts and historical reanalyses (Gichamo, Draper 2022). To represent the snowpack as correctly as possible, the available observations can be effectively combined with a background state generated by a numerical weather prediction model in the process of data assimilation. For snow observations, the statistical interpolation assimilation method is the state of the art used by most operational forecasting centres (Brown et al. 2003; de Rosnay et al. 2014; Renshaw et al. 2018; Birman, Mahfouf 2019).

At Czech Hydrometeorological Institute (CHMI), the ALARO configuration of the ALADIN NWP system (Termonia et al. 2018) is used for operational forecasting. This configuration is described in references (Brožková et al. 2024; Bučánek et al. 2015). Prior to the time of writing, the snow observations were not assimilated, and the snow representation was determined solely by the microphysics scheme and the snowpack parametrization described in Douville et al. (1995). To improve it, we make use of the available snow-depth observations by deploying and optimizing a statistical interpolation assimilation scheme.

## Data

For assimilation in numerical weather prediction models like ALADIN, the information about snow is usually derived from the in-situ measurements of snow depth and satellite observations. This contribution makes use of the former by combining data from the SYNOP station network with national snow observation networks from Czechia (CHMI), Austria (GeoSphere Austria) and Hungary (HungaroMet). In the Swiss Alps, additional data were provided by the Swiss WSL-Institut für Schnee- und Lawinenforschung SLF. In total, data from 2217 stations were gathered with about 50% of the stations originating from the SYNOP dataset. The remaining half of the stations makes the station network notably denser in central Europe, especially in the Alps.

The period selected for both the optimization and subsequent validation of the assimilation scheme is the winter season starting in December 2021, as this season was synoptically dynamic, featuring two heavy snowfall episodes in February and April.

## Method

The assimilation method used is the statistical interpolation algorithm, suitable for surface fields thanks to its relative simplicity and effectiveness. For the assimilation of snow depth observations, its single-variable and two-dimensional form is used. It can be described by the analysis equation (Daley 1991)

$$x_a = x_b + W[x_o - H(x_b)]. \qquad (1)$$

The vector of model background $x_b$ is combined with the increment of this background vector from the vector of observations $x_o$. To make these two vectors comparable, $x_b$ is transformed into the space of the observations by the observation operator $H$, accounting for both the spatial mapping between the model grid and observation locations and for the conversion from model prognostic variables, such as snow reservoir [kg/m$^2$] and snow density [kg/m$^3$] into the observed quantities like snow depth [m]. The weights given to the increment are determined by the weight matrix $W$. The form of this matrix can be derived from the requirement of minimum analysis error variance with the following result

$$W = BH^T(R + HBH^T)^{-1}. \qquad (2)$$

Here, the background and observation error covariance matrices $B$ and $R$ estimate the uncertainties of the model background and the observations and the correlations between them. As observation errors are considered uncorrelated, $R$ is a diagonal matrix with estimated observation error variances $\sigma_o^2$ on the diagonal. The errors of the forecast model are naturally correlated, and therefore the non-diagonal

elements of $B$ are modeled using various correlation functions. In our scheme, we tested two such functions, specifically a Gaussian and a so-called MESCAN correlation function. The main difference between these two functions lies in the treatment of the vertical correlation. While the former uses Gaussian decay with growing atmospheric pressure difference, the latter operates with a linear decline with increasing altitude difference.

Another essential part of the assimilation scheme is quality control, a multiround procedure checking the suitability of all available observations. It consists of a metadata check, optional spatial quality control, a so-called first-guess quality control, and a final round evaluating the results of the preceding ones. While spatial quality control tests spatial consistency of the observations, the first-guess round compares the modeled value to the observed one. In higher altitudes with more complex orography and more snow, these values can differ substantially. To account for this, an option making the benevolence of the first-guess round grow with altitude is available.

To test the different options of the snow analysis scheme, three different configurations were selected, tuned, and validated. The experiment **exp-FR** uses the Gaussian correlation function, no spatial quality control, and no first-guess quality control benevolence modification. Importantly, it rejects any observation made above 1500 m a.s.l. The configurations **exp-AT-gauss** and **exp-AT-mescan** both use the spatial quality control round and are more benevolent towards the observations from higher altitudes during the first-guess quality control procedure. Unlike **exp-FR**, they do not blacklist observations from above 1500 m a.s.l. They differ by the correlation function used.

All these three configurations have adjustable parameters which need to be tuned to suit the forecast model and the station network. As the assimilation scheme is highly interconnected, this tuning was approached as multidimensional optimization using the Nelder-Mead simplex algorithm (Nelder, Mead 1965). For n parameters to be optimized, this algorithm uses a geometric object with n+1 vertices, the simplex, which moves around the surface of a function to find its minimum. The function minimized is in our case a metric quantifying the error of the snow analysis, the weighted normalized mean square error of snow depth:

$$WNMSE = \Sigma_{i=1}^{4} \alpha_i \frac{MSE_i}{y_i^2}. \qquad (3)$$

Here, the summation is performed over four distinct altitude groups. The weighting coefficients $\alpha_i$ account for differences in station density, while the normalization by $y_i$ takes into consideration differences in mean observed values among these groups.

After convergence of the simplex algorithm, the most promising parameter settings were tested in longer experiments and examined from various points of view, such as the observation monitoring, parameter values analysis, the visual inspection of snowpack and analysis increments fields, and results of statistical scores calculated separately for different altitude groups. Based on these tests, the final parameter combinations for **exp-FR**, **exp-AT-gauss,** and **exp-AT-mescan** were selected and compared to each other for 54-hour forecasts during January 2022.

## Discussion

While all the optimized configurations yield results that are significantly better than the reference without snow analysis, there are clear differences between their performance, both in terms of the overall scores and the altitude-specific results. Concerning the RMSE, **exp-FR** provides the best results in the middle (500–1000 m a.s.l.) and mountain (1000–1500 m a.s.l.) altitude bands. However, as this setting ignores observations from higher altitudes, it fails to diminish the substantial BIAS, and consequently the RMSE above 1500 m when compared to the **exp-AT-gauss** and **exp-AT-mescan**. These two experiments show, to some extent, similar results, with the former excelling in altitudes above 1500 m and the latter performing better at all other altitudes.

From the point of quality control strictness, **exp-FR**, lacking the spatial quality control check, is the most lenient setting, followed by **exp-AT-mescan** and **exp-AT-gauss**. From this point of view, **exp-AT-gauss** is the most robust and error resistant setting.

Apart from the results of statistical scores and observation monitoring, the physical meaning of the parameter values, for instance of the horizontal length scale, was taken into consideration. While the experiments **exp-FR** and **exp-AT-mescan** use rather short horizontal correlation length scales of approximately 14 and 10 km, it is almost 19 km in case of **exp-AT-gauss**. While a shorter horizontal correlation length can be more appropriate in the mountainous regions, it can be less effective in parts of the domain with sparse station coverage, such as Ukraine or Belarus. The fact that **exp-AT-gauss** performs relatively well in higher altitudes while operating with larger horizontal correlations makes it a preferable option and, given its advantages outlined above, it has been selected as the final configuration for preoperational testing.

During such testing, the impact of the new assimilation scheme on other near surface variables was assessed, generally yielding mixed results. A significant surface cooling, affecting the screen-level temperature was observed, with a degradation in the forecast for shorter forecast ranges of up to 18 h and an improvement for longer lead times. Through saturation water vapour pressure feedback, the screen level relative humidity was also affected, with a similar pattern as in case of $T_{2m}$. Additionally, the forecast of geopotential was influenced through the air thickness feedback, with a positive impact on forecasts at all lead times.

Mitigation of the negative effects on screen level temperature and relative humidity was achieved by retuning the gridbox snow fraction. This quantity represents the part of the gridbox surface effectively covered by snow (Ševčík 2023) and impacts both the resulting gridbox albedo and emissivity. The result of this retuning can be seen in Fig. 1. The original experiment **exp-AT-gauss** without any snow fraction modification leads to clearly lower screen level temperatures than the snow-analysis free reference (black). The experiment with retuned snow fraction in green demonstrates the ability to mitigate the cooling, as it fits the reference BIAS, while keeping the performance of snow depth forecast unaffected.

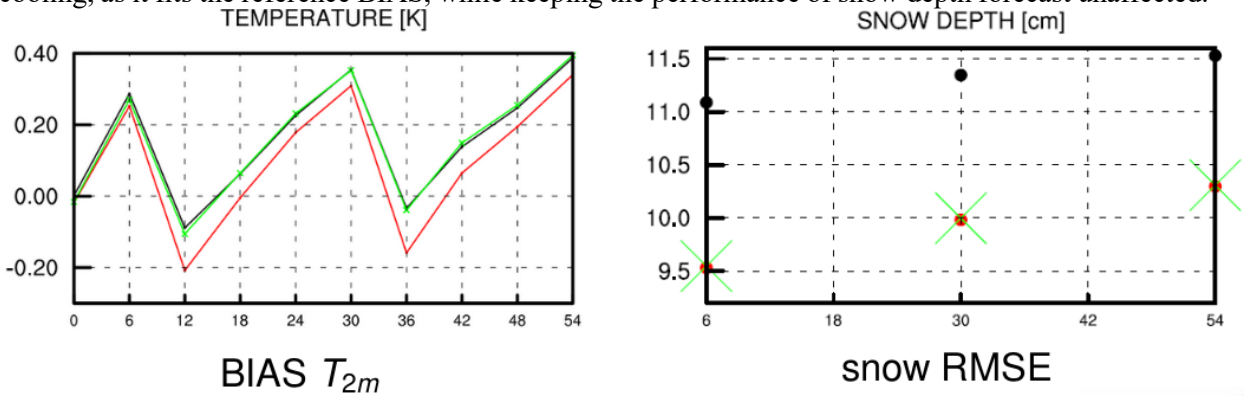



*Fig. 1. BIAS of screen level temperature and RMSE of snow depth forecasts for the verification period of 1–21 January 2026, snow-analysis-free reference (black), the **exp-AT-gauss** setting without (red) and with (green) retuned snow fraction.*

## Conclusion

The implementation of a snow depth assimilation scheme in the ALADIN numerical weather prediction model operated by the Czech Hydrometeorological Institute was described. The scheme uses the statistical interpolation method. Three different variants of the scheme were tested, to assess the impact of different correlation functions and modifications of quality control, especially in higher altitudes. The parameters of these configurations were tuned using the Nelder-Mead simplex algorithm, resulting in significant improvement of snow depth forecast when compared to the snow-analysis-free reference. After more detailed analysis, the experiment **exp-AT-gauss** was selected as the final configuration for operational use. As part of the preoperational testing, the effect of the snow analysis on other near surface variables was assessed and the negative effects were mitigated by adjusting the gridbox snow fraction.

At the time of writing, the scheme is a part of the operational suite of ALADIN at CHMI and could be potentially useful for the creation of historical reanalyses.

## Acknowledgments

The research was partially funded by the Technology Agency of the Czech Republic (grant no. SS02030040, PERUN).

## References

Barry, G. R. The Role of Snow and Ice in the Global Climate System: A Review. *Polar Geography*, **26**, 3, 2002, pp. 235–246. Retrieved from: https://www.researchgate.net/publication/249055062_The_Role_of_Snow_and_Ice_in_the_Global_Climate_System_A_Review

Birman, C., Mahfouf, J. F., 2019. Snow analysis for NWP at Meteo France and recent developments for land surface analysis [online] [cit. 5. 7. 2026]. Retrieved from: https://cimss.ssec.wisc.edu/iswg/meetings/2019/presentations/WEDNESDAY/1050_Birman.pdf

Brown, R. D., Brasnett, B., Robinson, D., 2003. Gridded North American monthly snow depth and snow water equivalent for GCM evaluation. *Atmosphere-Ocean.*, **41**, 1, pp. 1–14. https://doi.org/10.3137/ao.410101

Brožková, R., et al., 2024. Numerical Weather Prediction activities at CHMI [online] [cit. 3. 7. 2026]. Retrieved from: https://www.umr-cnrm.fr/accord/IMG/pdf/aws_2024_chmi.pdf

Bučánek, A., Trojáková, A. Brožková, R., 2015. Asimilační schéma BlendVar v ČHMÚ. *Meteorologické zprávy*, **68**, 5, pp. 129–136. https://www.chmi.cz/documents/42501/55105/chmu_mz_5-15.pdf/e836b86c-6145-af72-227f-ac1a5a152c69?t=1761301223405

Daley, R., 1991. Atmospheric Data Analysis. New York: Cambridge University Press. Second Edition. ISBN 0-521-38215-7.

De Rosnay, P., et al., 2014. Initialisation of land surface variables for numerical weather prediction. *Surveys in Geophysics*, **35**, pp. 607–621. https://doi.org/10.1007/s10712-012-9207-x

Douville, H., Royer, J., Mahfouf, J. F., 1995. A new snow parameterization for the Météo-France climate model. Part I: Validation in stand-alone experiments. Climate Dynamics, **12**, pp. 21–35. https://doi.org/10.1007/BF00208760

Gichamo, T. Z., Draper, C. S., 2022. An Optimal Interpolation–Based Snow Data Assimilation for NOAA's Unified Forecast System (UFS). *Weather and Forecasting*, **37**, pp. 2209–2221. https://doi.org/10.1175/waf-d-22-0061.1

Nelder, J. A., Mead, R., 1965. A simplex method for function minimization. *The Computer Journal* [online]. **7**, 4, pp. 308–313. Retrieved from: https://academic.oup.com/comjnl/article-abstract/7/4/308/354237

Renshaw, R., et al., 2018. Soil moisture and snow assimilation in a regional model at the Met Office. In: *Geophysical Research Abstracts* [online]. 2018. Retrieved from: https://meetingorganizer.copernicus.org/EGU2018/EGU2018-18173.pdf

Ševčík, J., 2023. Validation of snow cover forecast by numerical weather prediction model ALADIN. MFF UK. [online] [cit. 1. 7. 2026]. Retrieved from: https://dspace.cuni.cz/handle/20.500.11956/182553

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., et al., 2018. The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. *Geoscientific Model Development*, **11**, pp. 257–281. https://doi.org/10.5194/gmd-11-257-2018