# Deep learning-based multifrequency ground penetrating radar data merging

G. Roncoroni[1], E. Forte[1], I. Santin[1], and M. Pipan[1]

## ABSTRACT

Ground-penetrating radar systems with a single central frequency suffer limitations due to the unavoidable trade-off between resolution and penetration depth that multifrequency equipments can overcome. We develop a new semi-supervised multifrequency merging algorithm based on deep learning and specifically on bi-directional long-short term memory to automatically merge varying numbers of data sets collected at different frequencies. We implemented a new training strategy based only on the data set of interest, without synthetic or real training data sets. The proposed methodology is tested on synthetic and field data, to evaluate its performance and robustness. The merging algorithm can manage the complementarity of information at different central frequencies, properly merging different types of data. Results indicate not only a smooth transition in time, but, even more important, a remarkable broadening of the bandwidth thus increasing the overall resolution. Our approach is not limited to specific frequency components or geologic settings but can be potentially exploited to merge any type of data set with different spectral components.

## INTRODUCTION

Ground-penetrating radar (GPR) exploits electromagnetic (EM) waves to investigate the shallow subsurface. It is a versatile technique that can be adapted to different applications spanning from geologic, engineering, environmental, archaeological, glaciological, and forensic investigations, among others (Lombardi et al., 2022). The achievable resolution is a function of the physical characteristics of subsurface materials, being inversely proportional to the EM velocity and unavoidably decreases for increasing depths, as for all the geophysical techniques. Resolution can be improved by increasing the source frequency, i.e., the frequency content of the EM wavelet used for sounding. However, this leads to a reduced depth of investigation due to the stronger attenuation of high-frequency signals. A trade-off between the desired resolution and depth of investigation is therefore mandatory in all the GPR systems based on a single central frequency. Moreover, by using a single central frequency signal, as is usually done, imaging performance is limited, in particular when the subsurface geometry is complex and the targets have different dimensions (Lu et al., 2020). Modern multichannel (or multiarray) systems can provide multifrequency data sets but pose severe logistical constraints, especially when they operate on steep or rough terrains. In addition, issues arise in interpretation due to the large amount of data and their intrinsically different nature when collected with different central frequencies antennas.

To combine together at least two signals having different central frequencies, several algorithms and procedures have been proposed. Dougherty et al. (1994) were the first who suggested to exploit such a possibility. Since that time, several other methods have been implemented and tested. The final objective is referred as multifrequency compositing (e.g., Endres et al., 2004), multifrequency fusion (e.g., Xu et al., 2018, 2019), data fusion (e.g., Lu et al., 2020) or merging (e.g., Roncoroni et al., 2022). A comprehensive review of the state-of-the-art of the different proposed approaches and their peculiarities can be found in Alemdağ et al. (2022).

A detailed classification of the proposed methods to combine together data with different frequency components is out of the scope of this study, but, in general, there are two different strategies either based on 1D fusion (i.e., considering each A-scan — or portions of it — as separated entities), or based on 2D fusion (i.e., combining B-scans or portions of them as “images”, see e.g., Lu et al., 2020). The latter often are also referred to as image fusion (Zhao et al., 2021).

Here, for the first time, we propose a multifrequency merging algorithm based on semi-supervised bi-directional long-short term memory (LSTM) to automatically merge varying numbers of data sets with different central frequencies. Training and inference of the neural network (NN) are made on the same data set to be merged. In Roncoroni et al., 2022, merging of separated windows (i.e., partially overlapping portions of B-scans) of the peculiar RIMFAX Mars radar data set is presented; as an improvement of the previous approach, we here propose a general procedure that can be applied to merge any multifrequency data without constraints in terms of time window length and overlap, number and range of frequencies to be combined, sampling interval and number of samples.

The performance of the proposed algorithm is here discussed considering synthetic and field A-, B-, and C-scans (i.e., 1D, 2D, and 3D data, respectively). We demonstrate that the merged results are robust and do not require any amplitude balance or other specific processing steps. The bandwidth is expanded, and the resolution is increased, whereas the phases are not distorted, and no noise or spikes are introduced. Such a strategy may prevent possible misinterpretations, making easier data interpretation from large to small scales, in turn providing a more comprehensive understanding of the shallow subsurface. Merged results can be further processed, used for automated advanced interpretation techniques or attribute-based analyses, and exploited for data inversion.

## METHOD

The proposed approach is based on a new training procedure based on unsupervised learning. This allows us to use the same data set for training and inference: the kernel of the methodology is a bi-directional LSTM (see e.g., Schuster and Paliwal, 1997). The choice of this type of neuron is crucial for our goals. Bi-directional LSTM, represented in a simplified form in Figure 1, is a recurrent neuron with long memory that takes, as input, the time-dependent signal and its reversal in time versions. This neuron is pivotal for the results and performances of the methodology. Let's consider a simple signal, i.e., a Ricker wavelet, at $t = 10$ ns discretized with sampling interval $dt$

$$W(t) = \left[1 - \frac{1}{2}(\omega(t - dt))^2\right] e^{-\frac{1}{4}(\omega(t-dt))^2}. \tag{1}$$

With a normal LSTM neuron, we can get information from everything that happened before ($t$), but we cannot get any information on what is happening at $t > t + \Delta t$. Therefore, we need to consider the signal from the sides with respect to each sample to allow the NN to correctly understand the waveform, its changes, and how it moves over time.

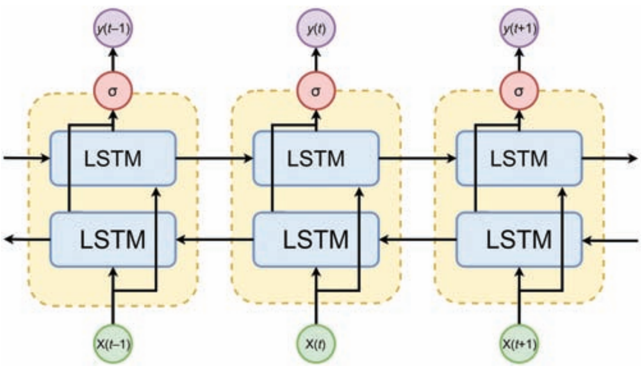

Figure 1. Unwrapped neuron of a bi-directional LSTM. We can see the recurrent nature of the neurons and the union between normal and reversed in time version with a sigmoid function.

After the definition of the neuron characteristics, we have to define a target for the training. Because we do not have a priori information on what we want to get and a synthetic simulation of the merged target could be tricky and not easy to define, we exploit a different strategy that mixes a supervised and an unsupervised methodology. We can define this approach as hybrid learning.

Before merging, we need to apply some essential preprocessing: one important step is the data amplitude recovery, usually exploiting exponential or more complex functions. This phase is crucial to correctly weight events when computing mean-squared errors. Another crucial step is to apply a normalization of the data, thus obtaining comparable amplitudes. We tried different approaches, and it turned out that the optimum solution was to scale the integral of the amplitude spectrum over the frequencies of interest to the same value, e.g., equal to one. Another possibility could be just to normalize the entire data set to one, but this would put excessive weight on the low-frequency components.

The proposed methodology is based on the training and inference of a NN, trained on the same data set to be merged. Specific limits and threshold can be inserted by the operator, e.g., the definition of weighting window for each frequency — as discussed in Roncoroni et al., 2022, but the procedure also can be applied to merge entire traces (i.e., A-scans) collected with different central frequency antennas. To merge data sets at different frequency bands, we used a procedure implemented for NN training, namely the gradient descent optimizations algorithms (Kingma and Ba, 2014), and the power given by a custom transformation based on only a few neuron weights.

The base of this approach is that merging should be performed by a few layers of a bi-directional LSTM (Schuster and Paliwal, 1997), 3 in this examples, with 4, 2, and 1 bi-directional layers and a single LSTM neuron for the output, aiming at minimizing the custom loss function with a single prediction from all input data. Exploiting only a few parameters for all data — as a result of using a very compact NN — we implemented a methodology that is robust to outliers and noise. A representation of this training scheme is depicted in Figure 2. Furthermore, the user can provide an input parameter, which we will refer to as the merging interval. This parameter ($\mathrm{val}_1$ and $\mathrm{val}_2$ in the following equations) is introduced into the algorithm directly in the loss function and should allow the NN to disregard areas where no useful input information is provided. If we want to maintain information inside all input signals, we can just set the full trace as the merging interval, obtaining in output a broader frequency spectrum signal

$$\mathrm{loss} = \lambda_1 \mathrm{loss}_1 + \lambda_2 \mathrm{loss}_2, \tag{2}$$

where $\mathrm{loss}_n$ is defined as

$$\mathrm{loss}_n = \|\mathrm{input}_{\mathrm{frequency}_n} - \mathrm{prediction}\|, \tag{3}$$

whereas $\lambda_n$ is defined as a piece-wise function:

$$\lambda_1(t) = \begin{cases} 1 & \text{if } t_1 < \mathrm{val}_1 \\ 0 & \text{if } t_0 \geq \mathrm{val}_1 \end{cases}. \tag{4}$$

$$\lambda_2(t) = \begin{cases} 1 & \text{if } t_1 \geq \text{val}_2 \\ 0 & \text{if } t_0 < \text{val}_2 \end{cases}. \tag{5}$$

With $\lambda_1$ and $\lambda_2$ referring to high- and low-frequency data, respectively. The introduction of the parameters $\text{val}_1$ and $\text{val}_2$, if $\text{val}_1 > \text{val}_2$, leads to the creation of three different zones, referring to Figure 2, within the green box, yellow box, and overlapping part; in detail:

- High-frequency zone (HFZ), where $t < \text{val}_1$, green box.
- Merging area (MA), where $\text{val}_2 < t < \text{val}_1$, overlapping area.
- Low-frequency zone (LFZ), where $t > \text{val}_2$, yellow box.

In the HFZ, the loss function is evaluated only on the highest frequency data. In MA, data are weighted with the same value, whereas in LFZ only low-frequency data are taken into account. After defining the two parameters $\text{val}_1$ and $\text{val}_2$, we train the NN with the same data we want to merge. The loss values as a function of training epochs is the measure of the convergence of the algorithm. So we have to consider these values to decide whether we have reached a stable model. After few iterations (see the "Discussion" section), a minimum in the loss function is reached and the same trained NN can be applied to infer the merged data.

## RESULTS

### Test on synthetic A-scan

As each data set is different, depending on the specific subsurface conditions being investigated, we may want different features to be enhanced and we let this specific option to the interpreter. A user-defined area is the kernel of the adopted strategy. Let us consider a double central frequency GPR signal, i.e., acquired with two different antennas, e.g., 400 MHz and 200 MHz. We may want to retrieve shallow information from the entire bandwidth, and we may want to maintain only the lower frequency components in the deepest section of the record; we therefore can define this as an information merging from now on.

If we are in a low loss medium, e.g., ice or dry sand, a higher central frequency can carry the most information, but adding lower frequency components can extend the bandwidth and improve resolution and information content. In Figure 3, we pointed out a typical simplified situation in which we can see HFZ (Figure 3a, gray), LFZ (Figure 3c, blue), and the merging area (Figure 3b, gray-blue). The synthetic data set was generated by convolving a maximum phase Ricker wavelet with just three reflection coefficients, and we set the training so that in the HFZ we just want to get the 400 MHz component (Figure 3a), in the LFZ just the 200 MHz one (Figure 3c), and in the merging areas (Figure 3b). As we can see, results are very good in terms of bandwidth widening of the merged data set, and the residuals are very low in the HFZ and LFZ for higher and lower frequency components, respectively. As we can see, the merging interval can retrieve a broader frequency version of the input wavelet, as apparent in the amplitude spectra (Figure 3b). Indeed, the residuals of the merged frequencies (Figure 3c) are almost perfectly symmetric and are null at the point at the peak of the wavelets, thus demonstrating the effective combination of the two frequency components, without any phase shift.

For the methodology, the phase consistency between input and output data is a crucial step. As we can see from the residuals, the phase is perfectly preserved because residuals have zero errors in these wavelets at the maximum peak.

### Test on synthetic C-scan

As a more complex and realistic synthetic application, we tested the methodology on the 3D synthetic Herten data set (Koyan and Tronicke, 2020a, 2020b), based on a publicly available hydrofacies data set, which has been derived from sedimentological outcrop mapping, direct sampling, and geostatistical simulations of an aquifer-analogue at the Herten gravel pit (SW-Germany). The model is very realistic because it includes several different sedimentary features, such as accretionary structures, cut-and-fill sequences, well sorted sand-gravel bodies, and matrix-supported gravels. The data set originates from six parallel digitized outcrop images, which are interpreted in terms of lithologic facies resulting in six 16 m long rasterized 2D facies sections. These sections cover a depth range of 7 m and with constant cell dimension equal to 0.05 m. Laboratory measurements of facies-specific hydrogeologic properties such as porosity and hydraulic conductivity lead to a subdivision of the mapped lithologic facies into 10 different units, termed hydrofacies. The authors use these 2D hydrofacies sections to perform geostatistical modeling resulting in a 3D subsurface model discretization of 0.05 m in all spatial dimensions, from which the electrical parameters are derived. Such a model was used as input to model 3D GPR data using gprMax (Warren et al., 2016, 2018). The Herten subsurface model was used to simulate a GPR common-offset data set with a 100 MHz central frequency source (Koyan and Tronicke, 2020a) then integrated with two other data sets obtained on the same subsurface model, but with central frequencies of 50 MHz and 200 MHz (Koyan and Tronicke, 2020b). The three synthetic GPR data sets are

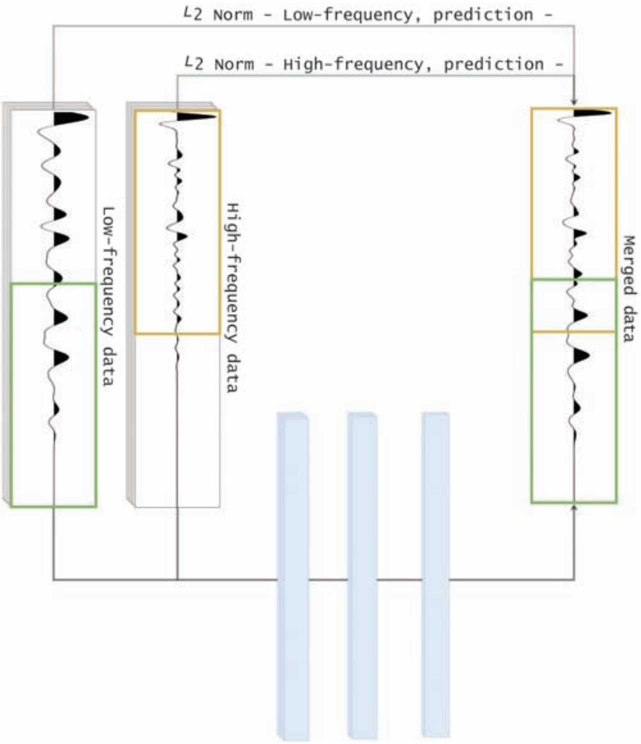

Figure 2. The NN training workflow. Starting from $n$ inputs (just two are here depicted), we retrieve a single mixed output. The network is trained to minimize the $L_2$ norm between each input and the output. A further mixing area constrain can be given to enhance higher frequencies in the shallower data.

perfect for testing the proposed merging procedure because, from one side, the data image different geometries and include diffractions and interference phenomena, and from the other the model parameter and geometries are perfectly known. All the details about model discretization and parameters used to derive the 50 MHz, 100 MHz, and 200 MHz data sets are reported in Koyan and Tronicke (2020a, 2020b).

In Figure 4, wiggle-variable area plots of just a few input traces (A-scans) and merged data are reported. We highlight not only the excellent phase preservation, but also the introduction of low-frequency components in the shallower part of the merged plot. For instance, if we compare the three wavelets of the event between 25 ns and 50 ns, we can clearly see the low- and high-frequency mixing in the merged data (Figure 4c).

In this test, we did not use the lowest frequency data (i.e., 50 MHz) because their information is very limited compared with the higher frequency ones (i.e., 100 and 200 MHz), with a significantly lower resolution and without any addition of information in the deeper part of the record.

The merged profile (i.e., B-scan; Figure 5c) shows not only enhanced overall resolution but also continuity of reflectors (see e.g., yellow arrows in Figure 5). The proposed method does not introduce neither coherent nor random noise and “cut and paste” effects (Roncoroni et al., 2022). Furthermore, we have a general broadening of the spectrum that can help inversion processes.

We applied the workflow to the entire 3D Herten data set (Figure 6) to further test the effectiveness and stability of the procedure.

The test proves the stability and robustness of the method and its effectiveness in preserving not only the signal amplitude contrasts, but also its phases. As we can see in 3D plot (Figure 6c), we obtained a good lateral continuity and reflectors matching at all the crossing points. Also, the imaging of relatively high-dipping reflectors in the merged data set is overall better than the one on the single central frequency volumes (e.g., yellow arrow in Figure 6).

## Test on field B-scan

The proposed merging method was finally applied to field GPR data. During the XXXVIII Italian Expedition to Antarctica (austral summer 2022–2023), a multifrequency GPR data set was acquired using a ProEx GPR instrument (Malå Geoscience) first connected with 250 MHz, then 500 MHz and finally 800 MHz shielded antennas. The data set is therefore composed of three GPR profiles collected on Boulder Clay glacier (Victoria Land — East Antarctica) along the same path with the three different antenna pairs. Details about the acquisition zone are reported in Azzaro et al. (2022). The profiles are collected along a straight line with an approximate eastwest direction, perpendicular to the glacier flow and with a 0.10 m trace interval. This data set was chosen to test the merging procedure because the different frequencies have a completely different penetration depth and highlight different glacial features such as diffractions and reflections with a large dip range.

The typical characteristics of GPR data sets, in which high central frequency antennas ensure high-resolution but have a limited penetration depth whereas the opposite is true for low frequency ones, is well highlighted by Figure 7, in which the 800 MHz GPR profile (Figure 7a) provides clear identification of reflectors in the first 60–70 ns. In particular, it images a strong continuous horizon very close to the surface, which can be associated to the transition between firn and ice. In 500 MHz data (Figure 7b), some internal glacial structures can be recognized, as well as the ice-bedrock high-amplitude reflection, imaged with higher continuity and toward its maximum depth (i.e., approximately 400 ns) in 250 MHz data (Figure 7c).

The corresponding amplitude spectra (bottom row of Figure 7) show the widened frequency band of the merged data. The comparison between single central frequency profiles (Figure 7a–7c) and the merged data (Figure 7d) highlights the overall improvement of resolution while ensuring a signal penetration depth, which is comparable with the 250 MHz one. In fact, all features characterizing each GPR profile can be recognized in the merged data, from the firn-ice transition to all internal high-dipping layers and the ice-bedrock interface. The 250 MHz antenna allows for optimal imaging of the bedrock, but the imaging

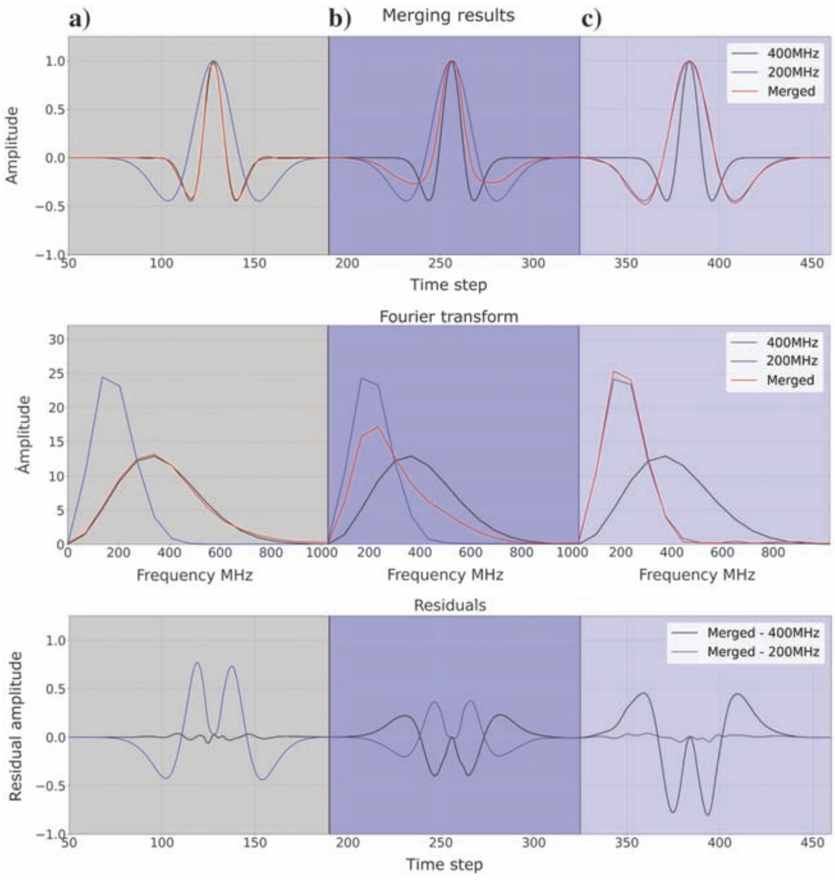

Figure 3. Frequency merging on a simple 3-reflectors synthetic model. The figure is divided into three columns: (a) HFZ, depicted in gray, (c) LFZ in blue, and (c) the merging area in gray-blue. In the first raw, we can see input signals and the merged result, in the second one the amplitude-frequency spectrum of each area and in the last raw the residuals.

offered by the merged data extends simultaneously from surface to deep targets ensuring for all of them optimal detectability and a very good resolution. In addition, low frequencies favour the identification of several dipping layers especially in the first 100 m of the profile, which were not clearly resolved from the 800 MHz and 500 MHz antennas.

## DISCUSSION

Most GPR data merging studies focus on synthetic data sets based on models that are very simple in terms of geometry and physical properties, i.e., planar reflectors with constant dip and constant EM properties of each layer (Xiao and Liu, 2016; Bi et al., 2020; Alemdağ et al., 2022). In some cases, random noise or spikes are added after forward modeling (e.g., Alemdağ et al., 2022). In contrast, we tested our procedure on one of the most realistic synthetic 3D GPR data set at present available and derived from a real outcrop through mapping, direct sampling, and geostatistical simulations (Koyan and Tronicke, 2020a, 2020b).

As far as real data are concerned, a typical problem is not only that low central frequency GPR data yield lower resolution, but the shallower portion is masked by the air and ground waves arrivals (Jol, 2009).

Figure 8 shows the same merged data as in Figure 7: the shallow unconformity (black arrows) cutting the dipping layers (red arrows), which onlap on it, can be easily identified especially in the first 100 m of the profile. At least 10 distinct internal layers, such as continuous and discontinuous horizons, also can be displayed and recognized, whereas in a single central frequency acquisition could not be possible to image such internal features. The amplitude spectra on the right of Figure 8 shows the spectral shift with depth due to the quick decay of high frequencies components but also the effectiveness of merging, which increased the resolution especially in the shallow portion of the profile, for which the bandwidth is remarkably wide and the peak amplitude is greater than 800 MHz. In the bottom part, the amplitude spectra are more constant and are characterized by a dominant frequency of approximately 250 MHz.

The stability of the method was tested by calculating some signal attributes and by migrating the merged data. The cosine of the instantaneous phase attribute (Figure 9) allows us to better evaluate the continuity of reflectors, and it is therefore optimal to assess the effectiveness and robustness of the merging method. In Figure 9, the ice-bedrock reflection can be traced clearly and continuously even where it has greater inclination (cyan arrows in Figure 9) and all the internal layers of the glacier. This example highlights the performance of the proposed method in preserving the continuity of the events without creating distortions or jumps related to the juxtaposition of different frequency bands.

Another test of the validity and efficiency of the merging methodology is the application of a migration algorithm to the merged data (Figure 10). Migration focuses the diffractions' energy to the apex and places the reflectors at their correct po-

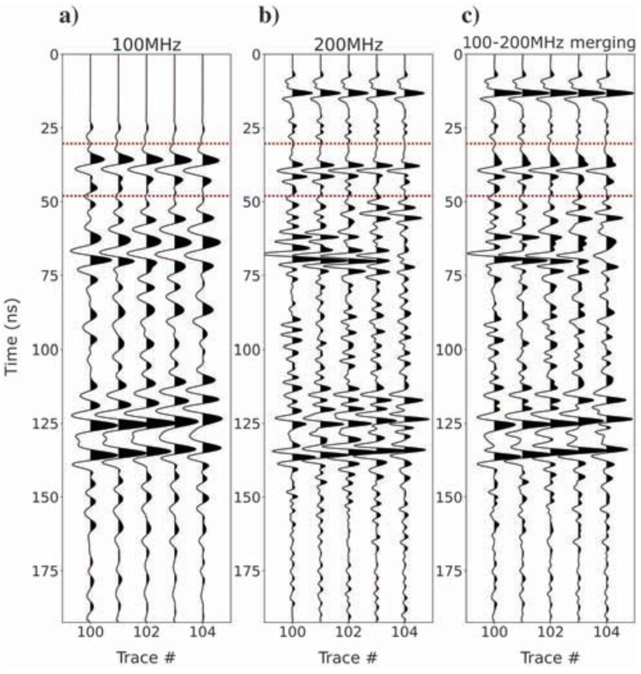

Figure 4. Merging test on five exemplary traces. (a) Wiggle plots of the 100 MHz, (b) 200 MHz inputs, and (c) the obtained merging. Red lines mark an area where the low- and high-frequency mixing is apparent.

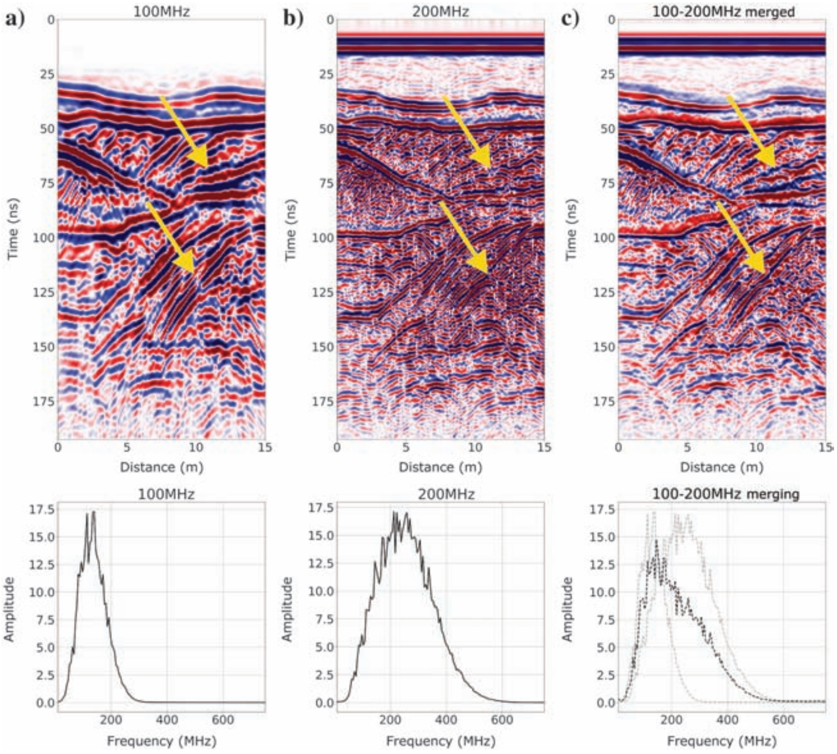

Figure 5. Merging test on a double-central frequency synthetic profile extracted from the 3D Herten data set. (a) 100 MHz profile, (b) 200 MHz profile, (c) merged profile, and their respective amplitude spectra. In the merged spectral panel, 100 MHz and 200 MHz spectra are reported in light gray for a better visual comparison with the merged one.

sition. All migration algorithms are indeed very sensitive to noise and abrupt spectral (in frequency and phase) variations of the signal (see e.g., Lehmann and Green, 2000).

After applying the topographic correction with a constant velocity equal to 0.17 m ns$^{-1}$, we applied a Stolt *f*-*k* migration algorithm with a constant velocity equal to 0.17 m ns$^{-1}$, which corresponds to a 3.2 relative electrical permittivity, typical of pure ice. This is surely an oversimplified velocity field, but it does not represent a crucial parameter for our test. The merged depth-migrated section, (Figure 10), provides the most realistic imaging of the internal structure of the glacier, allowing us to better appreciate the correlation between the morphology of the bedrock and the shape of the inter-

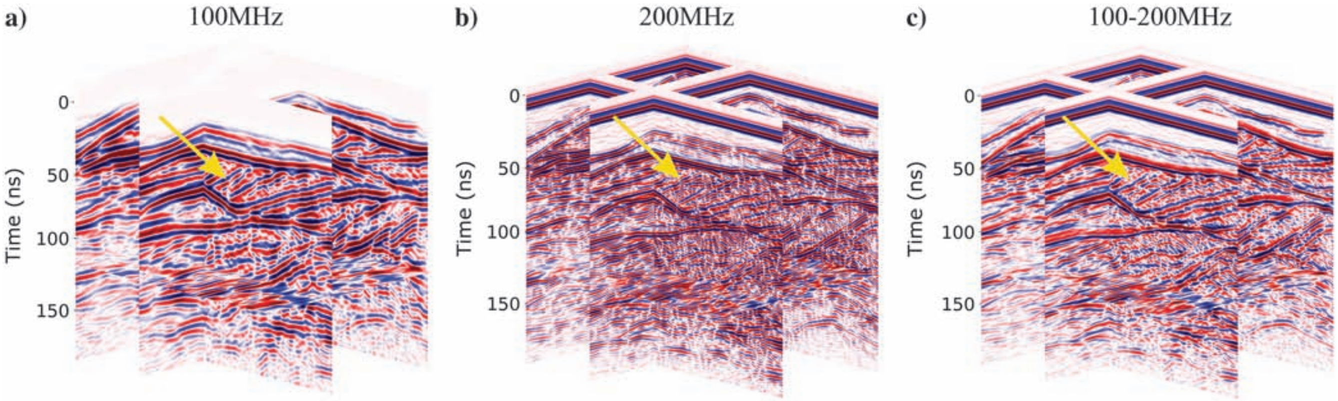

Figure 6. Merging of the entire 3D 100 MHz and 200 MHz Herten data sets: (a) 100 MHz, (b) 200 MHz, and (c) 100–200 MHz merged data.

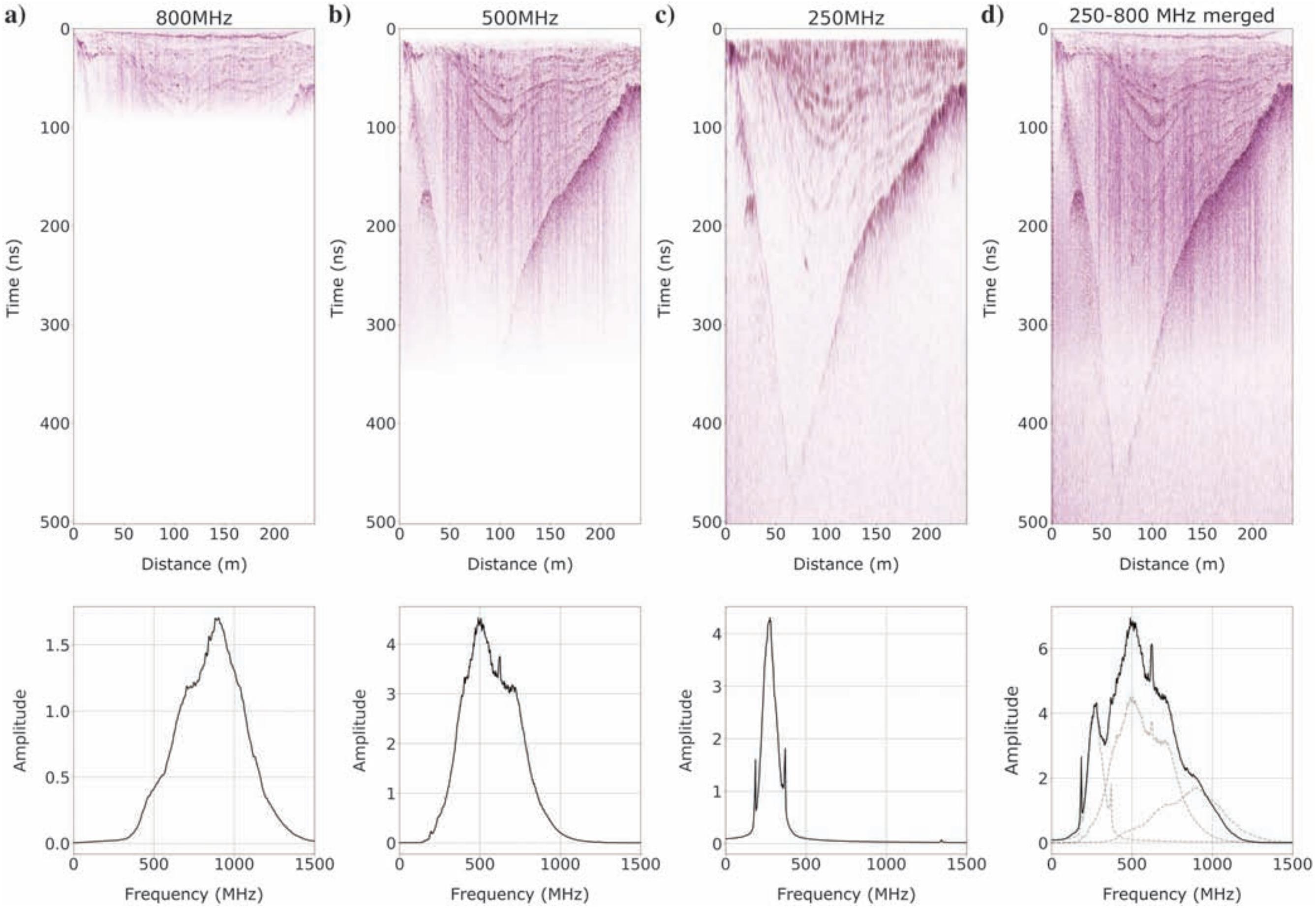

Figure 7. Merging test on a field data set: (a) 800 MHz profile, (b) 500 MHz profile, (c) 250 MHz profile, (d) merged profile, and their respective amplitude spectra. In the merged spectral panel, the 250 MHz, 500 MHz, and 800 MHz spectra are reported in light gray for a better visual comparison with the merged one.

nal layers, onlapping on the almost horizontal unconformity that cuts the deeper glacial strata. No apparent over or undermigration effects are present, whereas the lateral continuity of the reflectors is always preserved thus demonstrating once again the stability, robustness, and validity of the merging procedure. Merged data can be further processed and analyzed, considering that no artifacts, outliers, or coherent noise components are added by the data merging process.

Many of the existing merging algorithms need to be applied to data sets after intensive preprocessing. In fact, in several procedures, amplitude recovery (i.e., gain) and spectral balance are mandatory (Xiao and Liu, 2015; Coster and Lambot, 2018; Alemdağ et al., 2022). This introduces some subjectivity in the merged output and the sections obtained by combining different frequency profiles present abrupt transitions between areas dominated by different central frequencies. In addition, data fusion procedures often require nonphysical amplitude balance and preequalization, user-defined time window selections and postprocessing procedures to minimize unwanted cut and paste effects (Lu et al., 2020; Roncoroni et al., 2022). In contrast, our approach does not require any specific pre/postprocessing and can be applied even on raw data when the signal-to-noise ratio is high sufficient. In our procedure, the integral of the amplitude spectra is scaled to the same normalization value before data merging, thus obtaining a spectral balance and normalization at the same time. In addition, the proposed procedure does not have specific requirements on the frequency range

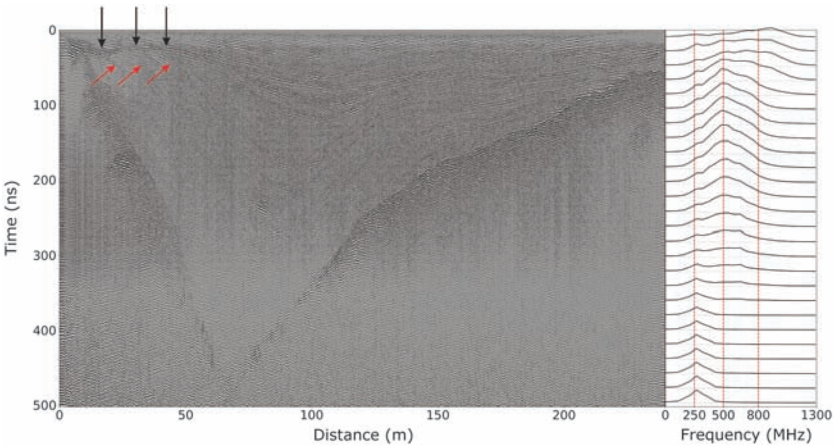

Figure 8. Merged data profile and amplitude spectra of the data in Figure 7d calculated with 20 ns long sliding time windows. The profile has a gradual frequency decay with depth down to approximately 300 ns. At greater depths the frequency content is more stable showing a dominant frequency of approximately 250 MHz. Red and black arrows mark specific reflectors described in the text.

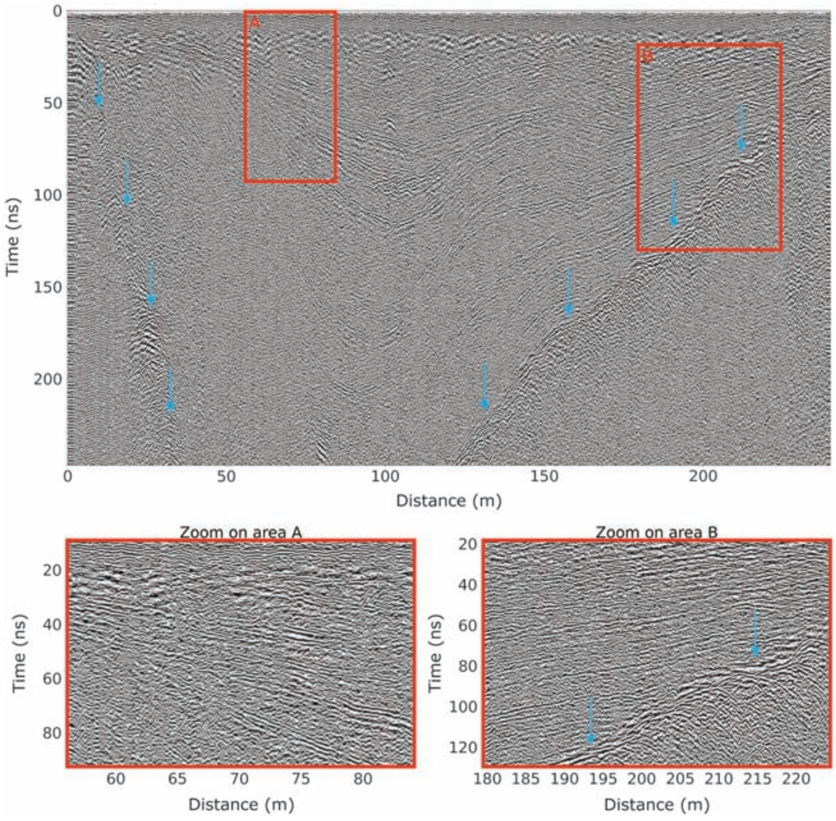

Figure 9. Cosine of the instantaneous phase calculated on the merged data of Figures 7d and 8. Time is limited to 250 ns with two close-ups on areas A and B, respectively.

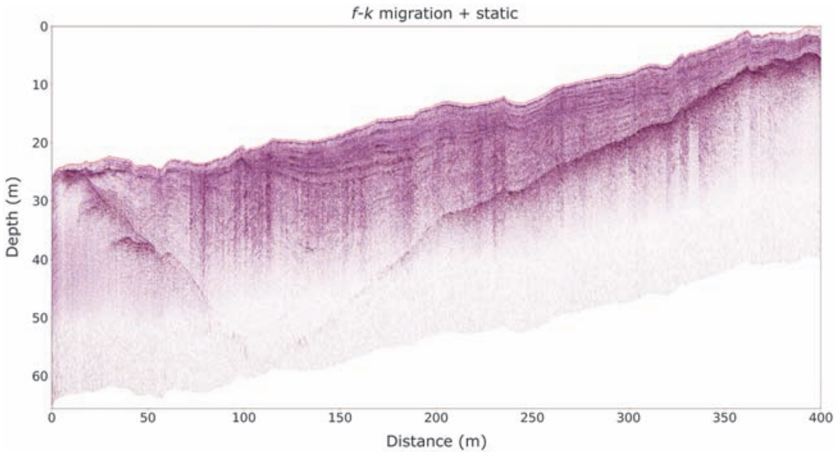

Figure 10. Depth-migrated merged field data as in Figures 7d and 8.

to be merged, nor to the number and sampling characteristics of the GPR profiles. The tests on real data (Figures 9 and 10) show that merged data can be further processed without any remarkable issue, in time and frequency domains.

The test of the proposed merging procedure on the field data set (i.e., the Antarctic glacier) took the following computation times on a laptop with Intel(R) Core (TM) i7-10875 H, 32Gb RAM and a Nvidia GeForce RTX 2070 Super with 8Gb of memory: training over 50% of each data set (i.e., every two traces) corresponding to 1200 by 4000 matrix took 1635 s, whereas merging (i.e., prediction) time less than 4 s. The trained NN, once calculated, can be applied to any data set with similar acquisition parameters. The computational costs are small enough to be managed by a laptop/PC.

## CONCLUSION

Merging GPR data with different central frequencies can optimize the unavoidable trade-off between resolution and penetration depth when using the standard approach exploiting just a single couple of antennas. We proposed a new merging algorithm that can effectively manage the complementarity of information at different central frequencies, properly exploiting data redundancy. We tested the method on different synthetic and real cases in totally different situations and with different acquisition parameters and frequency ranges. Stability tests demonstrated the robustness of the algorithm and the improved quality of the results, which can be further processed and analyzed. The proposed procedure is helpful not only to improve the response of reflectors but also of diffractors and in complex subsurface conditions, such as heterogeneous materials and conflicting dips. It can successfully manage even interfering events without introducing noise or artifacts. Another strength of the methodology is that it is totally data driven and can deal with noisy data. The training model requires a three order of magnitude longer computational time than the prediction. However, the training can be done just once and can then be used to merge any data set with similar spectral, geologic, and geometric characteristics. The proposed procedure is not limited to specific frequency components or geologic setting and can be potentially exploited to merge any type of data set having different spectral components. Further research will be focused on the application of this methodology to merging data coming from 3D GPR data sets. We would also focus on possible applications on common-shot gather stacking procedures, exploiting a modified version of this methodology.

## ACKNOWLEDGMENTS

The acquisition of Antarctic data are supported by a grant from the National Antarctic Research Program (PNRA), Research Project PNRA 2018_00186-E. Authors would like to thank drP. Koyan (University of Potsdam) for kindly giving us the 3D Herten synthetic GPR data set. Cineca consortium is further acknowledged for making available its high-performance computing facilities.

## DATA AND MATERIALS AVAILABILITY

B-Scan data are available at https://doi.org/10.6084/m9.figshare.23993688.v1. C-Scan data can be found at 10.17632/by3yh79hx4.1. Codes related to the methodology are presented in https://github.com/Giacomo-Roncoroni/merging_GPR.

## REFERENCES

Azzaro, M., M. Papale, C. Rizzo, E. Forte, D. Lenaz, M. Guglielmin, and A. Lo Giudice, 2022, Antarctic salt-cones: An oasis of microbial life? The example of Boulder Clay Glacier (Northern Victoria Land): Microorganisms, **10**, 1753, doi: 10.3390/microorganisms10091753.

Bi, W., Y. Zhao, R. Shen, B. Li, S. Hu, and S. Ge, 2020, Multi-frequency GPR data fusion and its application in NDT: NDT & E International, **115**, 102289, doi: 10.1016/j.ndteint.2020.102289.

Coster, A. D., and S. Lambot, 2018, Fusion of multifrequency GPR data freed from antenna effects: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **11**, 664–674, doi: 10.1109/JSTARS.2018.2790419.

Dougherty, M. E., P. Michaels, J. R. Pelton, and L. M. Liberty, 1994, Enhancement of ground penetrating radar data through signal processing: Symposium on the Application of Geophysics to Engineering and Environmental Problems, doi: 10.4133/1.2922053.

Endres, A. L., A. D. Booth, and T. Murray, 2004, Multi-frequency compositing of spatially coincident GPR datasets: Proceedings of the 10th International Conference on Ground Penetrating Radar, 271–274.

Jol, H. M., 2009, Ground penetrating radar: Theory and applications: Elsevier Science Ltd, 524.

Kingma, D. P., and J. L. Ba, 2014, Adam: A method for stochastic optimization: arXiv preprint, doi: 10.48550/arXiv.1412.6980.

Koyan, P., and J. Tronicke, 2020a, 3D modeling of ground-penetrating radar data across a realistic sedimentary model: Computers & Geosciences, **137**, 104422, doi: 10.1016/j.cageo.2020.104422.

Koyan, P., and J. Tronicke, 2020b, Analyzing 3D multi-frequency ground-penetrating radar (GPR) data simulated across a realistic sedimentary model: 18th International Conference on Ground Penetrating Radar, 275–278.

Lehmann, F., and A. G. Green, 2000, Topographic migration of georadar data: Implications for acquisition and processing: Geophysics, **65**, 836–848, doi: 10.1190/1.1444781.

Lombardi, F., F. Podd, and M. Solla, 2022, From its core to the Niche: Insights from GPR applications: Remote Sensing, **14**, 3033, doi: 10.3390/rs14133033.

Lu, G., W. Zhao, E. Forte, G. Tian, Y. Li, and M. Pipan, 2020, Multi-frequency and multi-attribute GPR data fusion based on 2-D wavelet transform: Measurement, **166**, 108243, doi: 10.1016/j.measurement.2020.108243.

Roncoroni, G., E. Forte, and M. Pipan, 2022, Merging gated frequency-modulated continuous-wave Mars2020 RIMFAX ground-penetrating radar data: Geophysics, **88**, no. 2, A7–A12, doi: 10.1190/geo2022-0466.1.

Schuster, M., and K. K. Paliwal, 1997, Bidirectional recurrent neural networks: IEEE Transactions on Signal Processing, **45**, 2673–2681, doi: 10.1109/78.650093.

Warren, C., A. Giannopoulos, and I. Giannakis, 2016, gprMax: Open source software to simulate electromagnetic wave propagation for ground penetrating radar: Computer Physics Communications, **209**, 163–170, doi: 10.1016/j.cpc.2016.08.020.

Warren, C., A. Giannopoulos, A. Gray, I. Giannakis, A. Patterson, L. Wetter, and A. Hamrah, 2018, A CUDA-based GPU engine for gprMax: Open source FDTD electromagnetic simulation software: Computer Physics Communications, **237**, 208–218, doi: 10.1016/j.cpc.2018.11.007.

Xiao, J., and L. Liu, 2015, Multi-frequency GPR signal fusion using forward and inverse S-transform for detecting railway subgrade defects: 8th International Workshop on Advanced Ground Penetrating Radar, doi: 10.1109/IWAGPR.2015.7292680.

Xiao, J., and L. Liu, 2016, Permafrost subgrade condition assessment using extrapolation by deterministic deconvolution on multifrequency GPR data acquired along the Qinghai-Tibet railway: IEEE Journal of Selected Topics in Applied Earth and Remote Sensing, **9**, 83–90, doi: 10.1109/JSTARS.2015.2487970.

Xu, X., G. Fang, J. Li, M. Liang, and Z. Li, 2018, Research on multi-frequency data fusion algorithm for ground penetrating radar: Progress in Geophysics, **33**, 2181–2186.

Xu, X., J. Li, X. Qiao, and G. Fang, 2019, Fusion of multiple time-domain GPR datasets of different center frequencies: Near Surface Geophysics, **17**, 141–150, doi: 10.1002/nsg.12033.

Zhao, W., L. Yuan, E. Forte, G. Lu, G. Tian, and M. Pipan, 2021, Multi-frequency GPR data fusion with genetic algorithms for archaeological prospection: Remote Sensing, **13**, 2804, doi: 10.3390/rs13142804.

Biographies and photographs of the authors are not available.