"MAGNETIC MODELLING AND ERROR ASSESSMENT IN ARCHAEOLOGICAL GEOPHYSICS: THE CASE STUDY OF URBS SALVIA, CENTRAL ITALY "

Annalisa Ghezzi[*,1], Antonio Schettino[1], Luca Tassi[1], Pietro Paolo Pierantoni[1]

[1] University of Camerino, School of Science and Technology – Geology Division, Camerino (MC), Italy

ABSTRACT

We perform an analysis of the errors that affect magnetic anomaly data in archaeological geophysics, arising from both survey time pro–cedures and common potential fields methods of magnetic data processing. Specifically, there are errors due to: 1. positioning of total field readings, 2. the estimated diurnal drift of the Earth's magnetic field, 3. the selected gridding algorithm, 4. the process of reduction of total field data to magnetic anomalies, 5. the application of decorrugation filters, and 6. knitting of two or more survey rectangles within the same archaeological area. Our analysis shows that in normal conditions these errors can have a magnitude up to few tens of nT and a lower limit exists to the amplitude of the anomalies that can be interpreted archaeologically. A correct error assessment is especially required when the magnetic anomalies must be interpreted quantitatively through a forward modelling procedure. We illustrate an application of these concepts to a magnetic data set acquired at the Roman settlement of Urbs Salvia (Central Italy). We show that forward modelling provides a powerful tool for the reconstruction of ancient buried settlements.

1. INTRODUCTION

The development of caesium–vapor magnetometers since the 1950s has found a fruitful application in archae-ological research, thanks to their high acquisition fre-quency and sensitivity. In fact, these features have allowed to perform accurate magnetic surveys, with the objective of measuring the tiny anomalies generated by buried ar-chaeological artifacts. At the same time, the diffusion of new generations of powerful computers and the design of specific software for the treatment of potential field data have further amplified the potentiality of one of the most common geophysical methods in archaeology.

In general, magnetic anomalies are fluctuations of the observed magnetic field intensity about a mathematical model of the Earth's main field, generated by induced or remnant magnetism of rocks or, at a smaller scale, by human artifacts. Induced magnetization results from the capability of some materials to be susceptible to presence of an external inducing field, while natural remnant mag-

netization (NRM) is a residual magnetization persisting for a long time period even in absence of external field. Usually, NRM of archaeological features is acquired as a consequence of burning, either during manufacturing or after fires.

Magnetic prospection represents one of the most useful and fast geophysical techniques for the detection of buried structures in archaeological sites. The magnetic method involves the measurement of the Earth's magnetic field intensity. Typically, total magnetic field intensities or vertical gradient are measured. The raw data are then subject to preprocessing and filtering to improve the signal quality before their transformation to digital raster images. In addition, in the case of total field data, the observed values are usually reduced to magnetic anomalies removing the Earth's core and crustal contributions. The magnetic maps obtained by these techniques are generally interpreted directly in term of archaeological features on the basis of the detection of specific patterns or alignments. However, the direct archaeological interpretation of magnetic maps doesn't allow establishing composition, geometry, and shape of buried artifacts, and the sources of the magnetic signal are almost always laterally displaced with respect to the corresponding anomalies. Finally, on the basis of a visual inspection of a magnetic anomaly map, it is generally hard determining whether a magnetic anomaly is generated by a single object or it results from the superposition of anomalies produced by neighboring objects.

In this paper, we will perform an examination of the different sources of errors in the production and interpretation of magnetic anomaly maps, accompanied by an analysis of the data uncertainty. We will consider the general aspects of the problem, in combination with a forward modelling approach to the determination of the pattern of buried structures at archaeological sites. Then, these concepts will be applied to the study of the magnetic data set collected at the ancient Roman city of Urbs Salvia, central Italy.

2. UNCERTAINTY AND ERROR SOURCES IN THE ACQUISITION AND ANALYSIS OF MAGNETIC DATA

Here we consider both the uncertainty associated with random acquisition errors and systematic errors that may occur during the subsequent processing phase. The latter cannot be considered as sources of uncertainty because do not arise from stocastic processes and cannot be treated statistically. Let $\Delta T^*(x,y,z,t) = T(x,y,z,t) - F(x,y,z,t)$ be the true anomaly generated by an archaeological feature at location (x,y,z) and time t, T and F being the observed total

field and reference field (core field plus crustal field) intensities, respectively. This quantity depends from time in two different ways. First, the geomagnetic field intensity is subject to the secular variation, which is of the order of few tens nT/yr and can be clearly observed at the scale of months. Apparently, this variation affects at the same time $T(x,y,z,t)$ and $F(x,y,z,t)$, thereby the two contributions could cancel out in the calculation of the anomaly ΔT^*. However, this conclusion would be wrong, because it can be easily shown that ΔT^* coincides with the projection of the anomalous field associated with archaeological artifacts, ΔF, onto the reference field direction $\hat{F}$ [e.g., Blakely, 1995]. This direction is subject as well to the secular variation, so that both inclination and declination of $\hat{F}$ change at the scale of months. In addition, the component of ΔF associated with induced magnetization is always parallel to $\hat{F}$, and changes with both direction and intensity of the geomagnetic field. Consequently, the magnetic anomalies are not time invariant and depend in a complex way from the secular variation of the main field. Another important component of the time variability of the true anomalies ΔT^* is associated with the external contributions to T, usually limited to solar–quiet diurnal variations. This variability is generally eliminated in the pre–processing phase of magnetic data treatment by a levelling procedure, which allows to estimate a diurnal drift curve $R(t)$ for the survey area and survey time.

An estimator of the true anomaly field ΔT^* will be referred to as an observed anomaly, $\Delta T(x,y,z)$ (for a specific epoch). This quantity will not depend from short–time field variations and in principle can be calculated applying the following expression to the acquired magnetic readings:

$$\Delta T(x,y,z) = f\left[T(x,y,z,t) - R(\mathrm{t}) - \sum_{n+m \leq N} a_n b_m x^n y^m\right] \quad (1)$$

where f is a filtering function, $R(t)$ is an estimated diurnal drift curve, and the polynomial coefficients a and b are estimated by a statistical regression of $T(x,y,z,t) – R(t)$ in such a way that ΔT has zero mean. Therefore, the validity of expression (1) is based on the assumption that: 1) the reference field F can be described adequately by a polynomial surface, and 2) the archaeological anomalies can be described by a random variable that represents deviations of the observed field from the reference polynomial surface. The anomaly field $\Delta T(x,y,z)$ calculated by expression (1) is affected by the following potential errors, which are discussed in detail below:

- Sensor positioning errors during the data acquisition;
- The variance of the observed total field values $T(x,y,z,t)$ about the estimated diurnal drift curve;

- Errors arising from gridding algorithms;
- Errors associated with a wrong choice of the polynomial degree N;
- Errors resulting from the application of the filter f;
- Errors associated with knitting of different grids from the same area.

2.1 UNCERTAINTY FROM POSITIONING ERRORS

An estimation of the uncertainty associated with positioning errors can be obtained starting from the maximum local variation of the observed total field, $T(x,y,z)$, with respect to an arbitrary unitary displacement from the current position, which is usually known as the analytic signal:

$$G(x,y,z) = |\nabla T| = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2} \quad (2)$$

To estimate the uncertainty associated with positioning, we need to provide an evaluation of the maximum displacement of the sensor around the theoretical position during a survey. The typical geometry of a magnetic survey is illustrated in Figure 1. The magnetic data are acquired at 10 Hz frequency along bi–directional survey lines equally spaced 0.5 m in a rectangular grid. At a typical operator velocity of 4 km h^{-1}, such a sampling frequency translates into an average 11 cm distance between readings. To evaluate the maximum displacement of the sensor around the theoretical position we need to consider the three components of sensor displacement around the recorded position during survey line walks (Figure 2), which depend on the specific terrain and environmental conditions as well as on the available equipment. In absence of a positioning control system [e.g., Bruniaux et al., 2018] the transversal component, ε_x, is generally of the
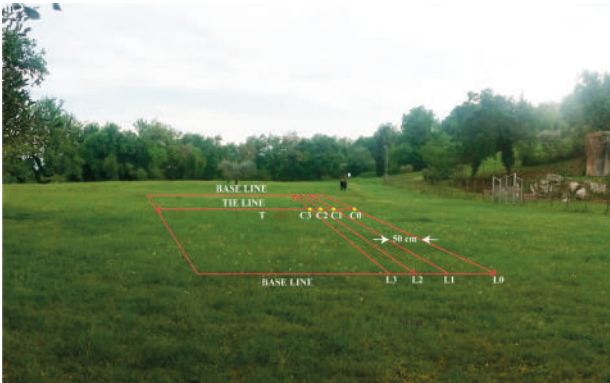

FIGURE 1. Typical mapped survey layout at sites with planar re–lief. Li (i = 0,1,...) and T are respectively the survey lines and the tie line (in red). Ci (i = 0,1,...) are crossover points (yellow dots) for levelling

FIGURE 2. Components of the maximum positioning error vec–tor during data acquisition.

order of 10 cm and is associated with a curved shape of the rope (e.g., during windy days) and/or small oscillations of the sensor about the rope. The longitudinal, ε_y, component results from various causes: a) magnetometer readings at fiducial marker points are not perfectly synchronized with pressing of marker commands; b) operator parallax displacements when pressing the fiducial command button, which can be systematically in advance or late; c) variations of velocity during the operator walk along a survey line, as well as from longitudinal oscillations of the sensor. Its magnitude is of the order of few tens cm and can be generally estimated taking the half–amplitude of zig–zag artifacts in the total field grid. Finally, a small vertical ε_z component is associated with irregularities of the terrain and does not exceed 5 cm in most cases. Assuming that $\varepsilon = (\varepsilon_x, \varepsilon_y, \varepsilon_z)$ is a random vector variable, the magnitude of the average position error, ξ, will be given by: $\xi = |\varepsilon|/2^{1/3} = (\varepsilon_x^2, \varepsilon_y^2, \varepsilon_z^2)^{1/2}/2^{1/3}$. We can estimate the local uncertainty, ε_P, associated with positioning errors by the following expression (Schettino et al., 2018):

$$\varepsilon_p(x,y) = \xi G(x,y) = (\varepsilon_x^2 + \varepsilon_y^2 + \varepsilon_z^2)^{1/2} G(x,y)/2^{1/3} \quad (3)$$

2.2 UNCERTAINTY FROM LEVELLING ERRORS

A magnetic survey is performed only when the value of the solar activity index Kp does not indicate magnetic storm conditions ($Kp < 5$). In our approach, a tie line T (Figure 1) is traveled once at the beginning of the survey in a short time interval after the initial time t_0 (~1–2 min).

Therefore, two successive readings are taken at the crossover points C_i of the intersections between survey lines L_i and T. The mismatches $T(x,y,z,t) - T(x,y,z,t_0)$ between two measurements at each crossover point are then used to build a diurnal variations curve $R(t)$. This function is estimated through a statistical regression of the crossover errors. The assumption of t_0 as a common initial time for all crossover points represents a key aspect of this approach and is based on the observation that the time necessary to walk the tie line T is negligible. The crossover errors, $T(x,y,z,t) - T(x,y,z,t_0)$, do not depend from the crossover point locations C_i, as they are determined exclusively from the variations of the geomagnetic field at each time t relative to the initial time t_0. The basic idea behind levelling is that these crossover errors form a time sequence that can be used to estimate the diurnal drift function through a statistical regression. An example of cubic polynomial regression of crossover errors is illustrated in Figure 3. We note the relevant dispersion, of the order of few tens nT, of the residuals about the estimated diurnal drift curve $R(t)$. If such dispersion could be interpreted as the result of rapid fluctuations of the geomagnetic field, we wouldn't have a new independent source of

uncertainty for the magnetic anomalies. Rather, we would eliminate the influence of the external field by generating a linear interpolant to the points (dashed line in Figure 3) and subtracting this curve from the observed total field data. However, geomagnetic micropulsations in the period range between 0.1 s and 10 min have amplitudes that rarely exceed 1 nT [e.g., Jacobs, 1970]. Therefore, the variance about $R(t)$ in Figure 3 must be considered as the result of random positioning errors, not as the result of a real physical process. This variance translates into an uncertainty of the time–independent total field $T - R$.

2.3 UNCERTAINTY ARISING FROM GRIDDING PROCEDURES

When dealing with scalar fields over geographic domains, it is common to represent the data by two–dimensional grids, formed by equally spaced nodes. The procedure of evaluation of a scalar field at grid nodes is called *gridding*. Data in grid format are suitable for a number of two–dimensional procedures, such as image processing and two–dimensional spatial filtering. When data are collected along lines that are roughly parallel, as in Figure 1, bi–directional gridding and kriging are ap-
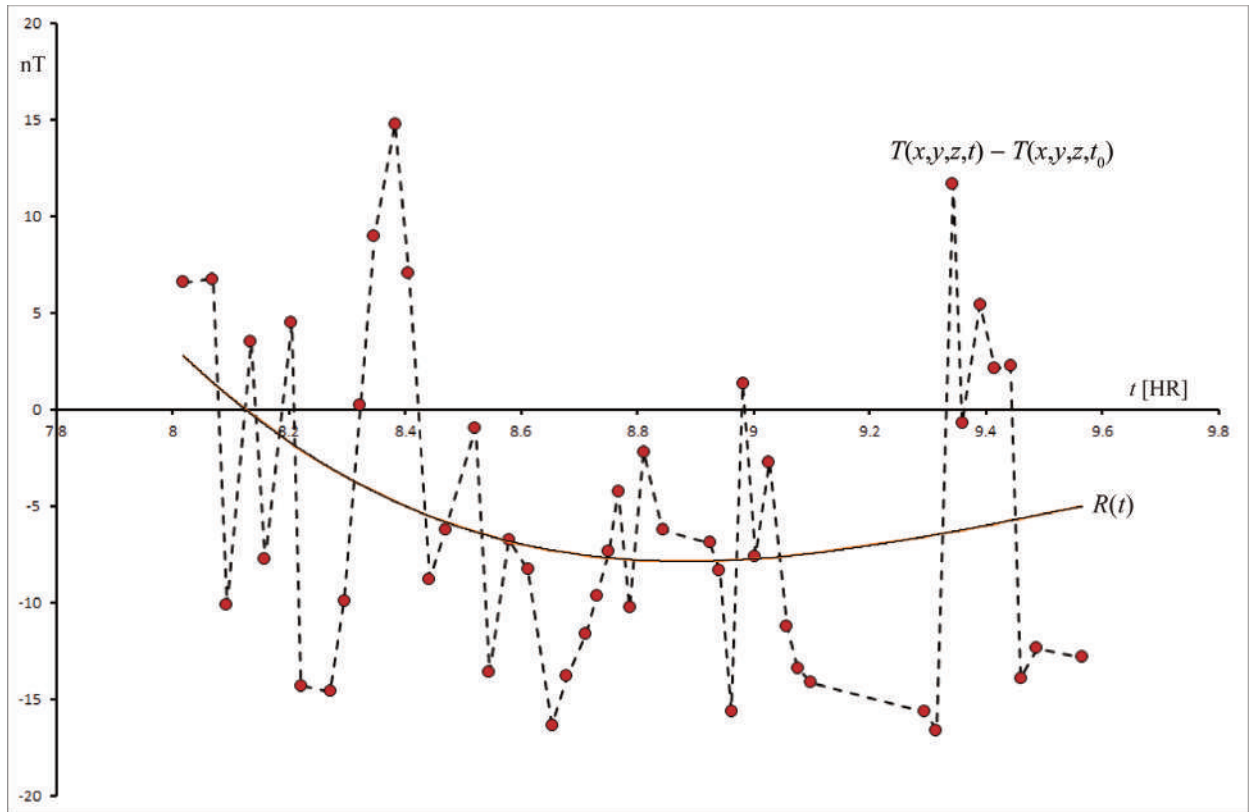

FIGURE 3. Example of diurnal drift curve $R(t)$ (brown line), obtained by cubic polynomial fitting of crossover errors $\varepsilon_i = T(x,y,z,t) - T(x,y,z,t_0)$ (red dots). In this survey, positions of magnetic data were measured using a GPS receiver configured to use corrections transmitted by a satellite–based augmentation system (SBAS).

propriate choices, especially if there is a high sampling density along the survey lines relative to the transverse direction. In particular, the bi–directional gridding algorithm [Smith and O'Connell, 2005] tends to emphasize trends perpendicular to the direction of the survey lines, in the attempt to fill the space between them. Alternatively, sparse data can be gridded using algorithms such as the Inverse Distance Weighted (IDW) interpolation and the Minimum Curvature algorithms. In general, only the kriging algorithm provides a built–in statistical estimation of the uncertainty associated with gridding. All the other algorithms require the application of independent validation procedures to obtain confidence bounds for the resulting grids. Different choices of the gridding algorithm may led to changes up to few nT on average, but the discrepancy may locally reach several hundreds nT (Figure 4).

The example in Figure 4 shows the variability of the results of gridding for four common algorithms through difference maps relative to the bi–directional procedure. These residual maps have means/standard deviations of 0.61 nT/14.62 nT (Fig. 4A), 0.12 nT/4.22 nT (Figure 4B), and 0.31 nT/7.96 nT (Figure 4C) respectively. Although the absolute magnitude of the average residuals is small, the standard deviation is comparable with the magnitude of many archaeological anomalies. In addition, the maps show a small circular zone where the residual reaches several hundreds nT. This could be a consequence of the Akima spline algorithm employed in the bi–directional gridding, which is less affected by sharp gradients.

Gridding algorithms can be considered as interpolation methods that introduce pseudo–random errors in the estimated field values at each grid location. Therefore, they represent an additional source of uncertainty. Such an uncertainty can be estimated directly in the case of the kriging algorithm [e.g., Bourges et al., 2012], because this is a geostatistical procedure that optimally predicts the field value at each grid node [Cressie, 1990]. Our experience suggests that this gridding method provides results very similar to those obtained using the bi–directional algorithm (cfr. Figure 4). For a typical survey, the kriging uncertainty is around ~1 nT, thereby we expect that an uncertainty of the same order of magnitude may affect bi–directional grids. More in general, the accuracy of a gridding algorithm can be assessed by cross–validation [Davis, 1987] and the confidence intervals can be estimated assuming that the residual between the original dataset and the pseudo–values follow Student's t distribution [e.g. Adisoma and Hester, 1996; Tomczak, 1998].

2.4 REDUCTION OF TOTAL FIELD DATA TO MAGNETIC ANOMALIES

The errors resulting from the representation of F by a polynomial surface in Equation (1) can be estimated as follows. It is assumed that the observed total field, T, includes three components with distinct wavelength bandwidths, respectively from archaeological, crustal, and core sources. More specifically, it is assumed that there is no or little intersection between the wavelength ranges of these three components of the total field. It is well known that the separation between crustal and core fields can be performed by removal of the IGRF spherical harmonic representation of the main field for the survey epoch [e.g., Schettino, 2014], which includes wavelengths above ~3000 km. Given that the greatest observable wavelength for a squared survey area of size L is $\lambda_{max} = L$, the core contribution is simply a constant $F(x,y,z) = F_0$ at the scale of archaeological survey areas (up to 100–200 m). Regarding the separation between crustal and archaeological anomalies, Schettino et al. [2018] have shown that the maximum wavelength of archaeological features is between 80 and 100 m and that there are no wavelengths in the observed magnetic field between this threshold and the maximum observable wavelength for a survey area of
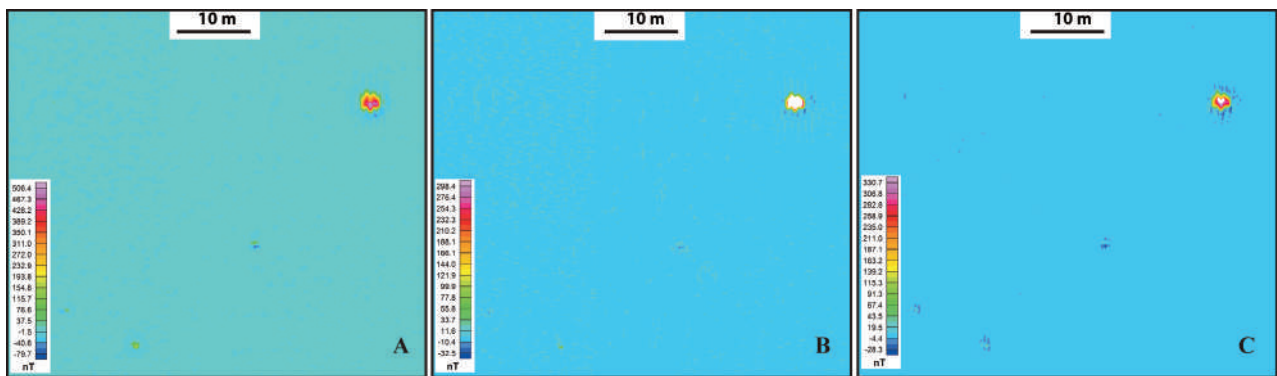

FIGURE 4. Errors associated with the choice of the gridding algorithm. The three maps show residual grids obtained subtracting from a bi–directional grid grids obtained by IDW (A), Kriging (B), and Minimum curvature (C).

size L, which is generally of the order of few hundreds m. Consequently, the reduction of total field data to archaeological anomalies by subtraction of a reference field F coincides with a HP filtering, where the cutoff wavelength λ_c depends from the polynomial degree N. This observation implies that the choice of N in Equation (1) is critical for obtaining an accurate representation of the archaeological anomalies. Schettino et al. [2018] noted that in many situations a value $N = 1$ is the correct choice for the polynomial representation of the reference field in areas not exceeding 50 m. Larger areas, up to 200 m width, could require a value of N between 4 and 5. Figure 5 shows that the error arising from an incorrect choice of the polynomial degree may locally reach tens nT, although the standard deviation over the entire grid will generally keep below 10 nT. Again, this value is comparable with the amplitude of most archaeological anomalies.

2.5 ERRORS FROM KNITTING PROCEDURES

The assembly of magnetic data sets from individual surveys often involves merging of two or more adjacent gridded data sets into a single composite grid. Although grid pre–processing procedures imply the removal of possible systematic noise, adjacent magnetic anomaly grids will not generally match along their common edges, so that a simple mosaic of the gridded data is often not recommended to draw an interpretable anomaly map [Cheesman et al., 1998]. Therefore, a correction is necessary along the grid edges to minimize the misfit between border data. In general, there are two possible, very different, situations. First, we may have a set of non–overlapping grids that must be simply joined in a composite map. Unfortunately, commercial computer programs for the treatment of potential field data (e.g., Oasis Montaj™) do not provide an algorithm that performs smoothed stitching of two or more
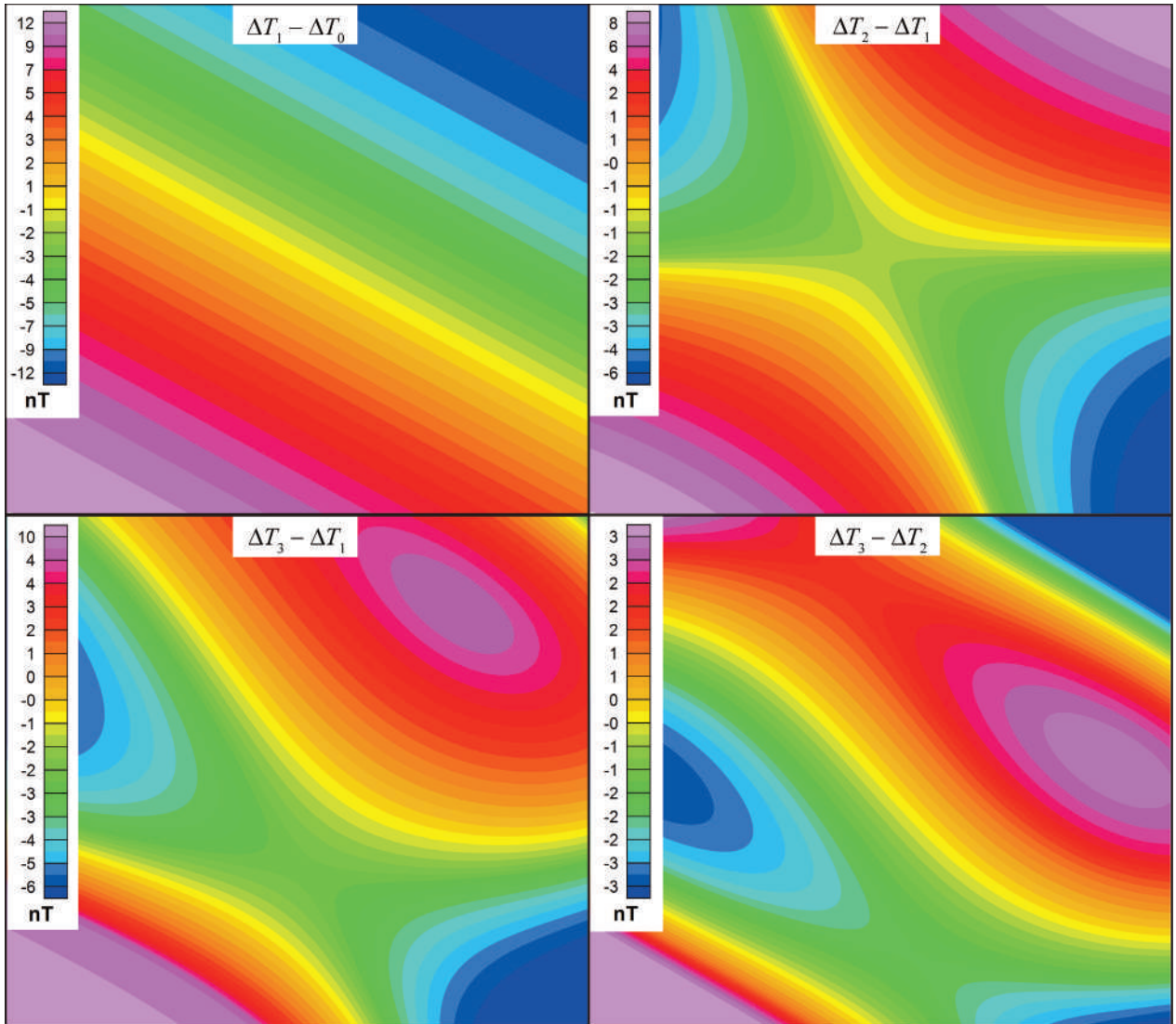

FIGURE 5. The Residual grids obtained subtracting anomaly fields that have been calculated using different degrees N for the poly–nomial representation of the reference field. ΔT_k indicates that the magnetic anomalies have been calculated setting $N = k$.

non–overlapping grids. Therefore, to date this situation will generally lead to the production of maps characterized by discontinuities along the former borders of the individual component grids (Figure 6).

In the case of overlapping edges, the misfit involves both long and short wavelength components of the signal. The primary goal of this kind of knitting is to perform corrections at different wavelengths with a minimum of distortion of the original data. Methods for merging two grids depend on the definition of a suture path in the overlap region R, and the construction of two subsets, R_1 and R_2, that include R and possibly part of the non–overlapping areas of the two grids, where data will be modified to produce a smooth continuous function along the suture line. This method uses Fourier analysis to decompose the error function along the suture path into a sum of sine–type functions with different spatial wavelengths. Corrections to data in R_1 and R_2 are applied independently to each wavelength and then summed, in order to obtain a smooth transition to the area outside R. More specifically, the size of these regions is not fixed, but determined dynamically for each wavelength, so that corrections are never applied beyond one quarter of the current wavelength from the suture line. The result will be a nearly seamless grid that eliminates the discontinuities along the suture line.

An alternative *blending* method distributes the corrections over the area of overlap on the basis of a weighted average of the two data values, taking into account of the minimum distance of each point in R from the two grid borders. The difference between the two methods is illustrated in the example of Figure 7, which shows the residual between two composite anomaly grids, obtained using the suture and blending methods. We note that the two techniques give similar results, the difference between them being of the order of few nT. Note that the transition zone in Figure 7 extends beyond the overlap area because its width is related to the presence of components of relatively long wavelength in the error curve along the suture path of the two adjacent grids.

2.6 DECORRUGATION FILTERS

An important step in the processing of raw total field magnetic data consists into the removal of some short–wavelenght artifacts (zig–zags) associated with small errors in the positioning of magnetic readings along the survey lines. To eliminate these artifacts, people generally apply the following basic procedure, although more sophisticated algorithms exist [Fedi and Florio, 2003 and references therein]: 1. High–pass filtering of the raw total field data using a high–order Butterworth filter (e.g., with $n = 8$) and a cutoff frequency depending from the corrugation wavelenght; 2. Filtering of the residual grid by a n–degree directional cosine in the profile direction; 3. Subtraction of the resulting grid from the original raw anomalies. Of course, this procedure modifies the amplitudes of the observed field to some extent, thereby it can be considered as an additional source of processing errors.
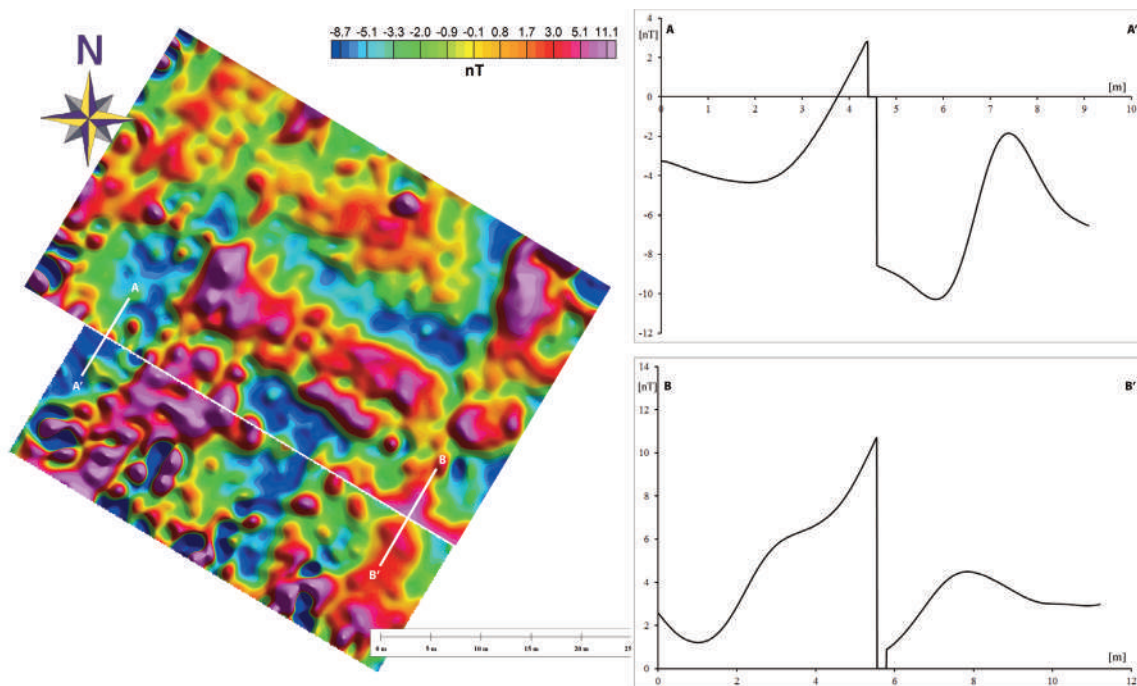

FIGURE 6. Knitting of two grids without overlapping edges. The two profiles reveal an important discontinuity across the two areas.
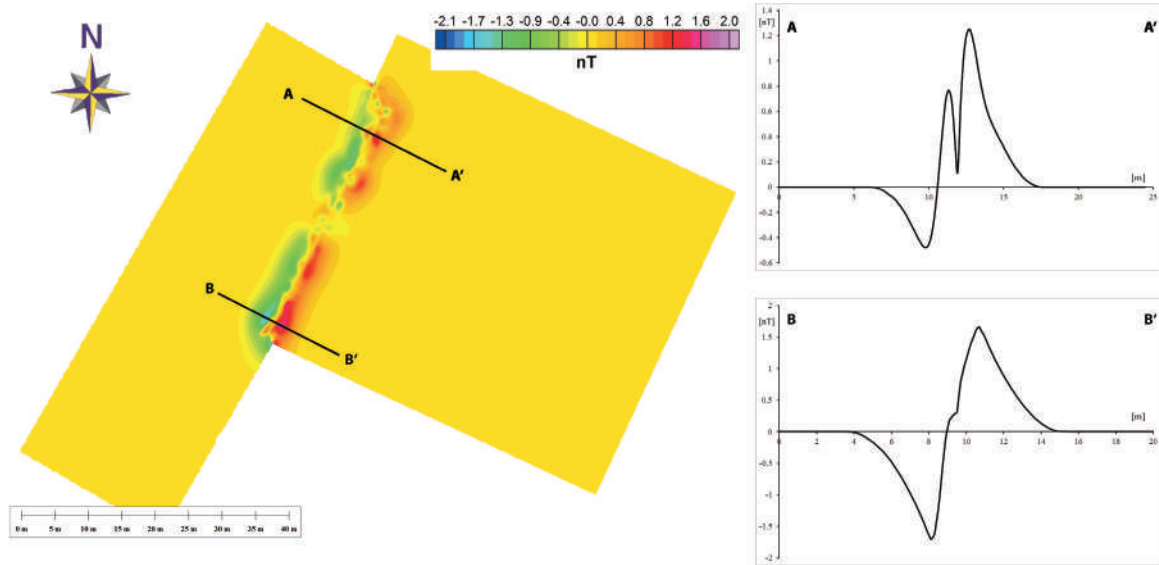
FIGURE 7. Residual map obtained subtracting a grid created using the suture method from one created with the blending method. Two profiles are drawn, which show only small (< 2 nT) differences in the area of overlap.

However, the amplitude of the corrections to the raw data is generally very small (a few nT) and cannot be represented by a random variable. Rather, this kind of filtering helps the "randomization" of the total field anomalies about the reference field.

3. FORWARD MODELLING OF MAGNETIC ANOMALIES

Three common techniques can be applied in the quantitative interpretation of archaeological magnetic anomalies: 1. Filtering of the observed data, which can enhance specific features [e.g., the edges of structures, see Nabighian et al., 2005 and references therein]; 2. Euler's deconvolution, which belongs to a class of techniques for the "automatic" magnetic sources depth estimation [Thompson, 1982; Desvignes et al., 1999]; 3. Forward and inverse modelling of the magnetic anomaly field. In this paper, we focus on the latter class of techniques, in particular on the application of new techniques of forward modelling in the archaeological context.

Forward modelling of magnetic anomalies is a trial–and–error procedure in which the interpreter operates interactively with a specific software to improve the fit between the synthetic anomalies generated by a magnetization model and the observed anomalies. The user starts from a hypothetical initial magnetization model and changes progressively the physical and geometrical parameters of the magnetic sources until the misfit falls below an accepted level that can be either defined arbitrarily or calculated rigorously according to an uncer-

tainty analysis. Although several general–purpose computer programs have been designed for the forward modelling of potential field data [e.g., Caratori Tontini, 2012], none of these codes allows an easy modelling of archaeological features. In addition, these computer programs do not consider at all the problems related to data uncertainty discussed above.

A new interactive forward modelling software, *ArchaeoMag*, has been designed at the University of Camerino for the specific needs of magnetic data modelling in archaeological geophysics [Schettino et al., 2018]. It can be freely downloaded at: http://www.serg.unicam.it/Downloads.htm. Differently from other general purpose forward modelling programmers [e.g., *for_3DFFT_mag*, Caratori Tontini, 2012] *ArchaeoMag* is written in C++ and does not rely on external runtime libraries such as MatLab®. In addition, it allows to distinguish between induced and NRM components of magnetization, thereby allowing a fine calibration of the model and possibly a dating of firing events. Three basic shapes and one composite object can be created using the *ArchaeoMag* GUI: Dipoles, rectangular prisms, general vertical prisms, and stairways. Each object can have specific magnetization parameters, size, and burial depth. The shapes can be easily edited, moved, rotated, or resized according to a trial–and–error procedure to obtain a better fit of the model anomalies to the observed values. To this purpose, *ArchaeoMag* allows to create anomaly profiles along user–defined traces, which show the synthetic and observed anomalies, as well as the misfit curve and the data uncertainty for checking the feasibility of the current model (Figure 8).
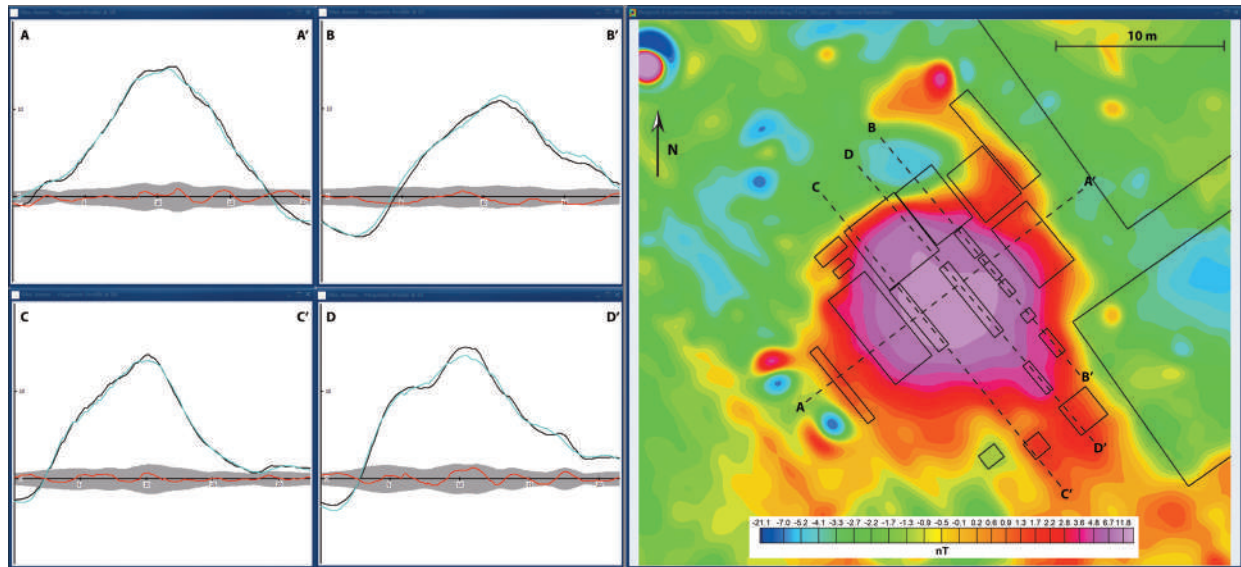
FIGURE 8. Example of modelling of coalescent archaeological anomalies at the Roman settlement of Hadrianopolis, southern Alba–nia. The magnetic profiles show observed and model anomalies (black and aqua lines, respectively), and the misfit curve (observed – calculated, in red), along selected traces. The grey area represents the total uncertainty associated with posi–tioning, processing, and instrumental errors.

The example in Figure 8 illustrates the potentiality of forward modelling methods. Magnetic modelling reveals that in reality this anomaly results from the superposition of several anomalies produced by neighbour vertical prisms with different NRM parameters. This is also evident on the basis of a visual inspection of the four magnetic profiles, which show several undulations corresponding to distinct contributions of objects having different magnetizations or parts of the same archaeological feature with different di-rections or intensity of magnetization. In general, both mag-netic dipoles and regular vertical prisms with homogeneous magnetization generate anomalies with two or three extreme points and a regular trend along any profile, depending on their magnetization parameters and the ambient field. There-fore, the presence of several extreme points in a magnetic profile or in its horizontal derivative are indicative of the presence of coalescent anomalies. In the example illustrated in Figure 8, most likely associated with a large Roman fur-nace, the coalescent anomalies result, at least in part, from variations in the NRM within the furnace rather than from distinct objects [Powell et al., 2002; Aidona et al., 2008].

In most situations, the direct visual interpretation of mag-netic anomalies leads to an erroneous reconstruction of the location and shape of buried structures, in addition to the impossibility of obtaining information about the physical parameters of the magnetic sources. However, it must be emphasized that even the most accurate magnetization model will be affected to some extent by the characteristic non–uniqueness of potential field data.

In general, there are two sources of ambiguity in the modelling of magnetic anomalies: 1) Infinitely many distri-butions of magnetization can reproduce the observed signal at the degree of accuracy required by the uncertainty distri-bution, and 2) there are special distributions of magnetiza-tion, called annihilators [Parker, 1977], that generate an anomalous field $\Delta F = 0$. It can be shown that annihilators are layers of constant thickness, draped on topography [Parker and Huestis, 1974], whose special distribution of magnetization produces a null anomaly field. Consequently, in the archaeological context only the topsoil could poten-tially assume this role, although it can hardly happen that the distribution of magnetic susceptibility in this layer matches the theoretical distribution that annihilates the anomaly field. In general, the common source of non–uniqueness in archaeological geophysics is associated with data uncertainty.

4. A CASE STUDY: THE ANCIENT ROMAN CITY OF URBS SALVIA

We now illustrate the concepts discussed above through a case study. Magnetic data were acquired in 2015 and 2016 at the ancient Roman city of Urbs Salvia, lo-cated in central Italy (Figure 9). The main objective of this survey was to reconstruct the urban organization of the city forum for determining possible sites of future ex-cavations. We found a complex pattern of buried struc-tures, possibly resulting from the coexistence of republican and imperial artifacts and burned structures.
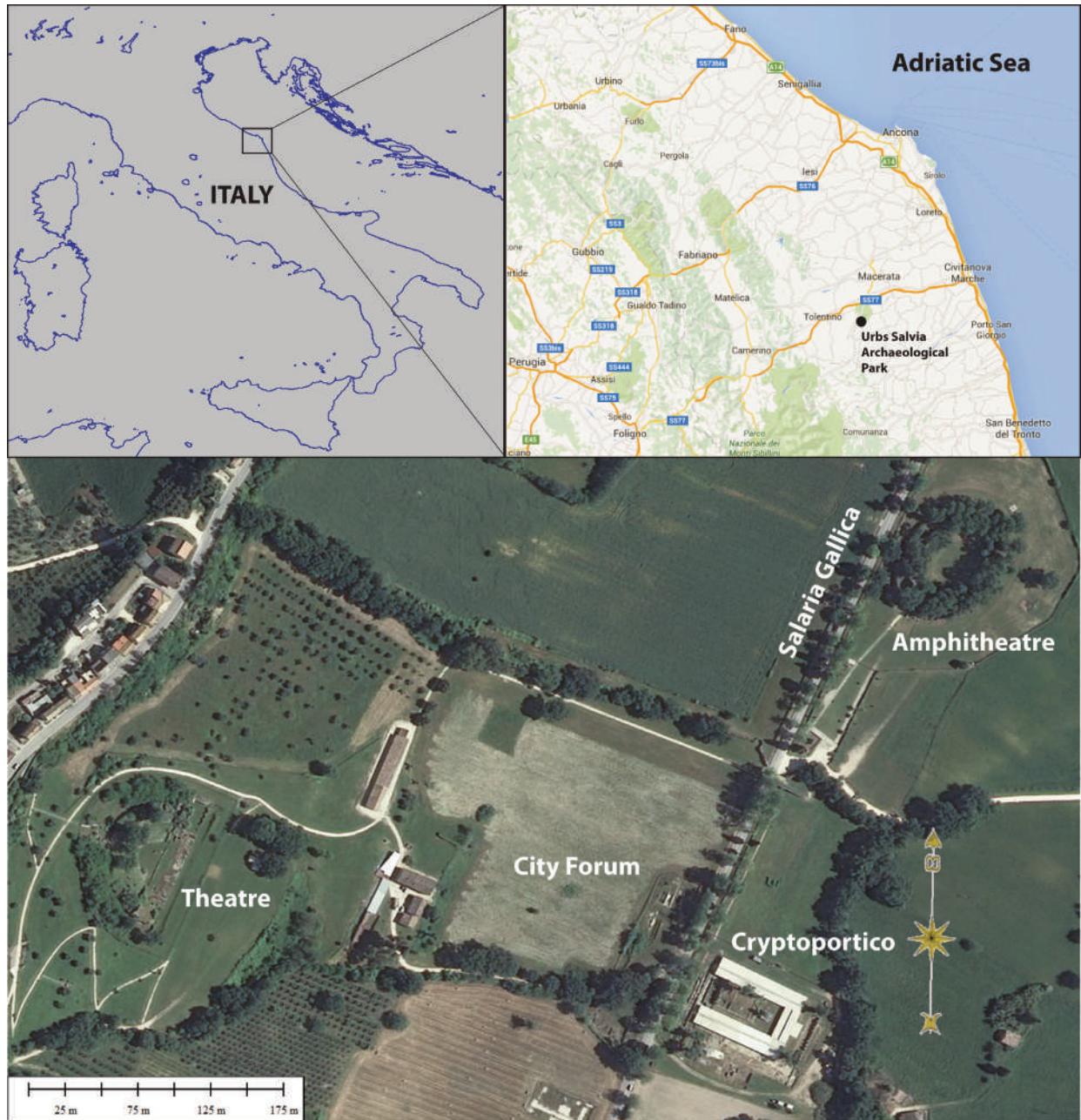
FIGURE 9. Location of the Roman settlement of Urbs Salvia.

Total field magnetic data were collected on a terrain clearance in a flat area using a Geometrics G–858 caesium vapor magnetometer. The survey area was divided in 14 rectangles having dimensions between 30x30 and 50x50 m², which were assembled together at the end of the processing and rotated into an UTM geographic reference frame. The magnetic data were read using the technique illustrated in Figure 1 at 10 Hz frequency (that correspond to an average 11 cm distance between readings) along bi–directional survey lines equally spaced 0.5 m. All the total field measurements were performed in solar–quiet conditions, with Kp index not exceeding 2. Finally, the data were corrected for the daily variations of the geomagnetic field through the levelling procedure described above and underwent standard pre–processing consisting into despiking and decorrugation procedures.

To reduce the total field readings to magnetic anomalies, we used the comparative procedure proposed by Schettino et al. [2018]. This procedure allows to select rigorously the polynomial degree N in Equation 1. We obtained a value of $N = 4$ for the composite grid, which included all the 14 rectangles, thereby the total field data were reduced to magnetic anomalies by subtraction of the corresponding low–degree best–fitting polynomial. The complete magnetic anomaly map of the Urbs Salvia forum is shown in Figure 10. This map shows the typical regu-
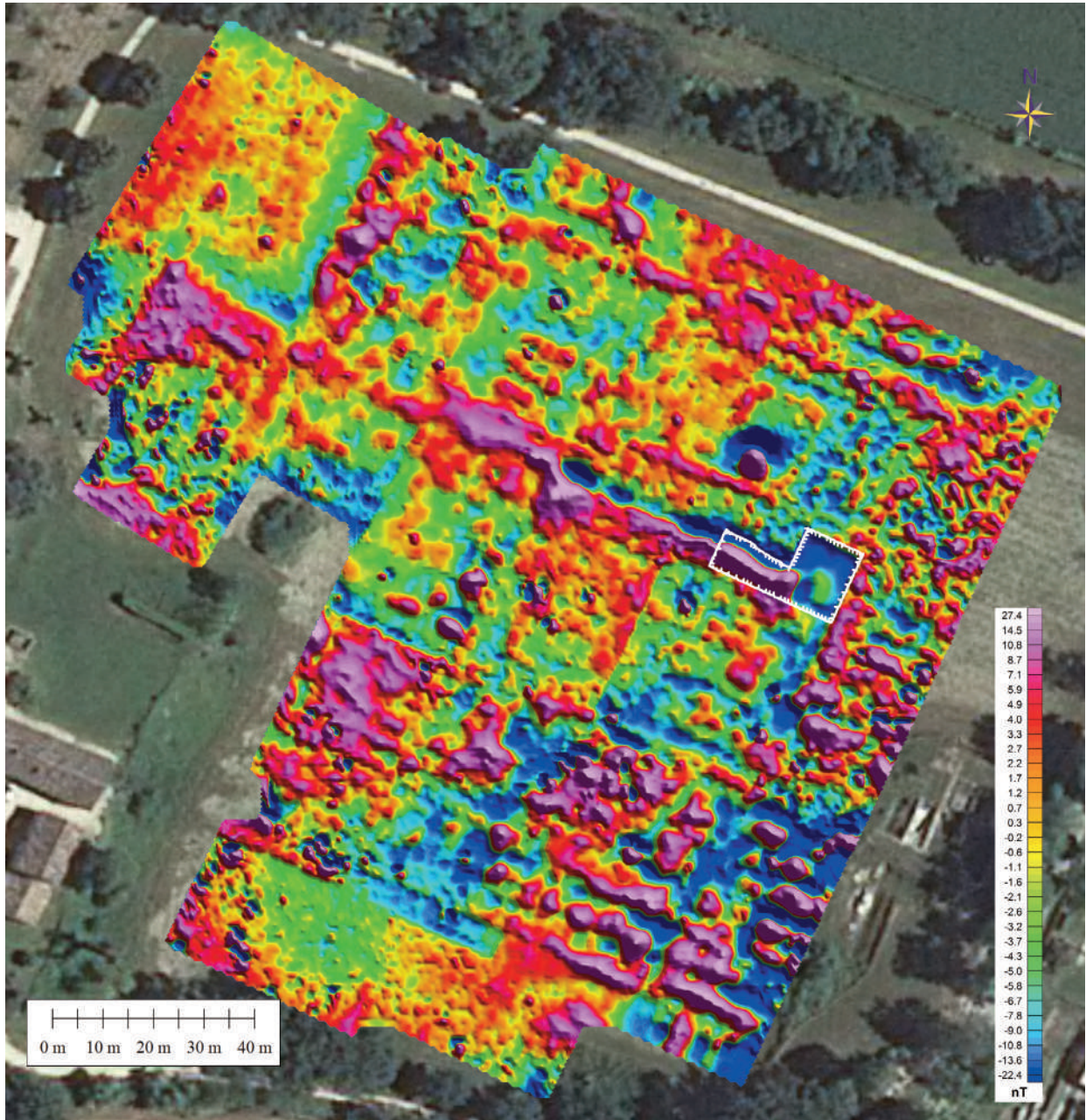
FIGURE 10. Magnetic anomaly map of Urbs Salvia. The white box shows a partially excavated area, where topsoil had been removed.

lar arrangement of Roman structures. The white box shows the location of a partially excavated area, where 1 m topsoil layer had been removed. Here the survey was performed at 25 cm resolution and the resulting magnetic anomalies were upward continued by 1 m to allow a correct knitting with the adjacent areas. However, in the subsequent modelling step it was anyway necessary to take into account of the absence of the topsoil layer. An artifact of the map shown in Figure 10 is represented by some sutures between adjacent rectangles, which do not correspond to real discontinuities of the field. These features arise from the mosaic knitting procedure that was applied to non–overlapping areas. Although the data in Figure 10

can be interpreted visually, this approach will not always allow a correct reconstruction of the archaeological features that are potentially responsible for the anomaly field.

Figure 11 shows the total uncertainty associated with both positioning errors of the magnetic data and the reduction to magnetic anomalies through Equation 1. While the former is proportional to the total field gradient, the latter represents a small "background" uncertainty associated with the statistical fit of the total field data to a reference polynomial surface. In the case of the magnetic data acquired at Urbs Salvia it resulted: $\varepsilon(x,y) = 0.1485|\nabla T| + 0.2209$ nT, where $\varepsilon_0 = 0.2209$ nT is the background uncertainty. The magnetization model illustrated below was
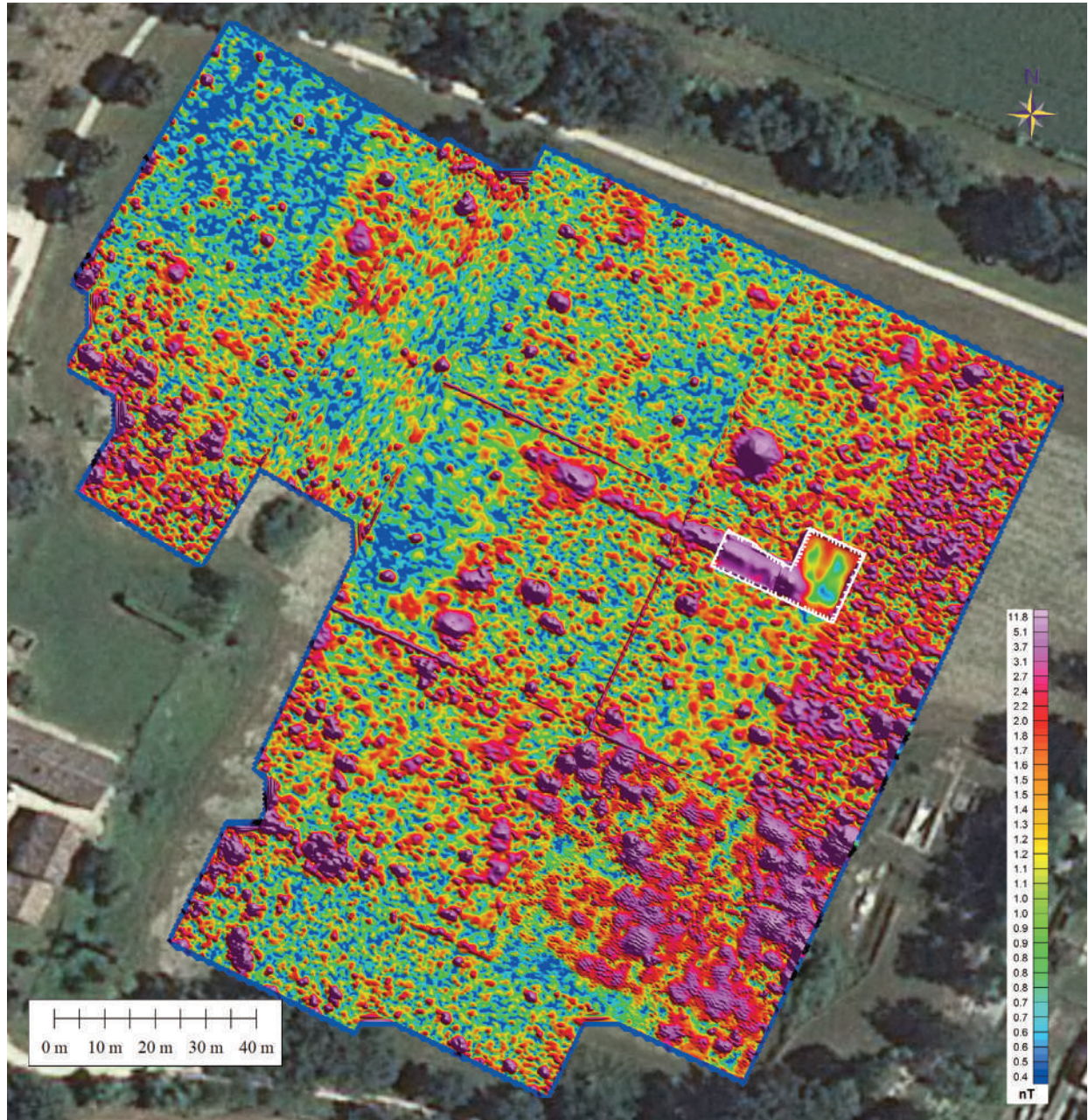
FIGURE 11. Uncertainty grid obtained using Equation 3 and maximum positioning error components ε_x = 0.10 m, ε_y = 0.18 m, and ε_z = 0.05 m. The y component was estimated on the basis of zig–zag amplitudes of raw total field data. The grid has a standard devi– ation of 4.90 nT.

built assuming a statistical ensemble of sources with aver- age burial depths determined by the radially averaged power spectrum illustrated in Figure 12 [Spector and Grant, 1970]. In some cases the burial depths were modi- fied during the modelling procedure to improve the fit. The plot in Figure 12 shows the presence of two main shallow contributions to the power spectrum. The average depths to the top of the two source distributions were calculated by the following equation [Spector and Grant, 1970]:

$$z_{top} = -\frac{s}{4\pi} \qquad (4)$$

where s is the slope of a linear tract of the power spec- trum function. We obtained z_{top} = 1.86 m and z_{top} = 1.03 m for the deep and shallow distributions of magnetic sources, respectively. These results show that reduction of the total field data to magnetic anomalies through a fourth-degree polynomial effectively removed the deep crustal (geological) and core components from the mag- netic signal, leaving an anomalous field representative of archaeological sources.

We performed a separate very–high resolution mag- netic survey (25 cm line spacing) of the partially exca- vated area after removal of the top 100 cm soil. During
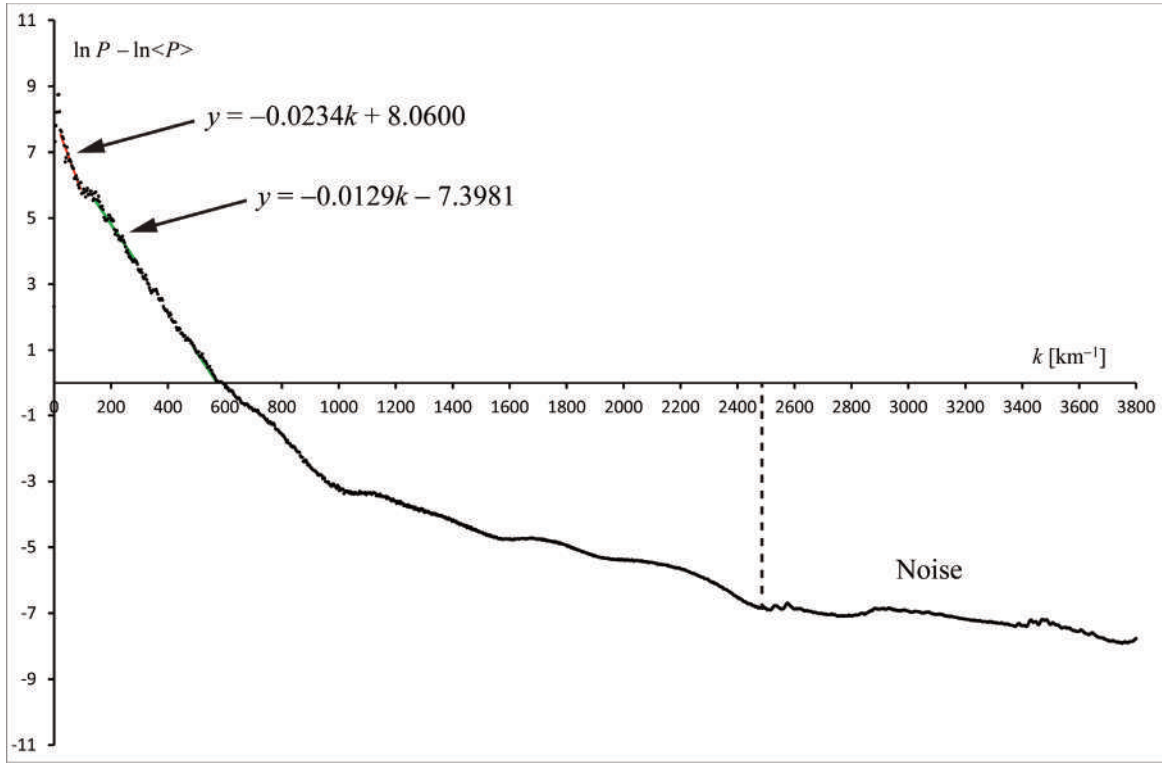
FIGURE 12. Radially averaged power spectrum of the Urbs Salvia forum magnetic anomaly grid (see Fig. 10), normalized by subtrac–tion of the log of the average spectrum density. Deep sources are represented by the dashed red fitting line. Shallow sources con–tribute to the range $140 \leq k \leq 580$ km^{-1} (dashed green line).

the subsequent modelling step, we compensated for the absence of the top layer of soil within this area introduc-ing an imaginary body having the shape and volume of the removed soil and zero magnetic susceptibility. In gen-eral, topsoil of settlements has a higher susceptibility rel-atively to the normal soil. This susceptibility, χ_0, must be subtracted from the susceptibility χ of an incapsulated human artifact in the calculation of the induced magne-tization M_I:

$$M_I = -\frac{\chi \text{-} \chi_0}{\mu_0} \, F \equiv \frac{\Delta\chi}{\mu_0} \, F \qquad (5)$$

where F is the regional geomagnetic field, μ_0 is the mag-netic permeability in the vacuum: $\mu_0 = 4\pi \text{x} 10^{-7}$ H/m, and the volume susceptibilities are expressed in SI units. To consider this component of magnetization in the compu-tation of the total anomalous field, we calculate the total magnetization vector, M, of a buried structure by the fol-lowing expression:

$$M = M_I + M_R \qquad (6)$$

where M_R is the remnant magnetization vector. In the case of the partially excavated area shown in Figure 10, we defined an air body having $\chi = 0$ and filling the vol-ume of soil removed by the excavation. According to

Equation (5), this body will have anyway an induced magnetization $M_I = -(\chi_0/\mu_0)F$ that contributes to the total anomalous field.

To illustrate the use of forward modelling techniques in combination with an uncertainty assessment, we can consider an interesting group of anomalies that cross the partially excavated area mentioned above (see Figure 10). These magnetic data include a long WNW–ESE lin-eament that results from the coalescence of several anomalies associated with smaller features at the level of confidence imposed by the uncertainty analysis, whose magnetization model is shown in Figure 13. In this instance, we wish to highlight that our approach al-lowed us to discard different models, for example the presence of a unique long wall based on a visual inter-pretation of the anomalies. The subsequent completation of excavation confirmed the existence of a complex pat-tern of structures related to the collapse of a Roman bath. Conversely, the initial archaeological hypothesis of a long and large wall, based on the visual identification of the long linear anomaly peak, was proved to be false. In conclusion, an accurate forward modelling procedure ac-companied by error assessment can be a powerful tool for the correct analysis of magnetic anomaly data and the identification of buried structures.
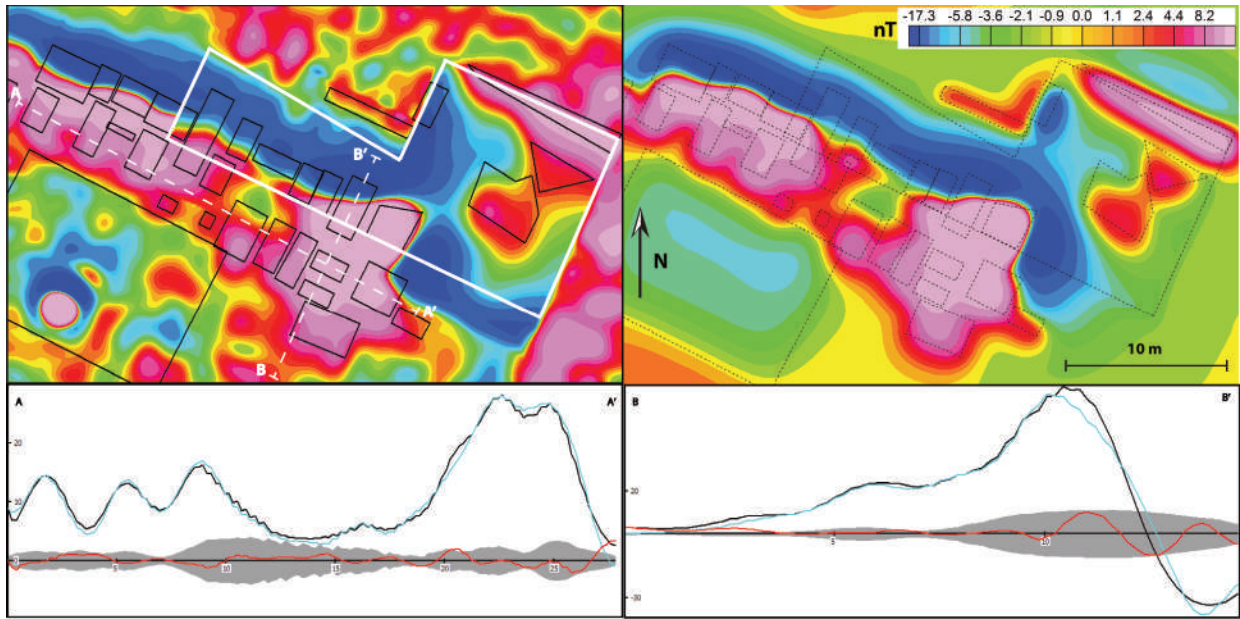
FIGURE 13. Modelling of an area close to the excavation (see Fig. 10). Upper left panel shows the observed anomalies and a model of the underground that explains most of the observed signal in this rectangle (black lines). A–A′ and B–B′ are the traces of the two profiles shown in the lower part of the Figure. Upper right panel shows a map of the theoretical anomalies, calculated from the block model. The lower panels illustrate the fit between observed and calculated anomalies along two profiles. The pale blue and black lines show model and observed anomalies, respectively. The red line shows the error curve, obtained subtracting calculated anoma–lies from the corresponding observed values. The grey areas represent the estimated uncertainty ε_P about the observed curve. The interactive procedure of modelling has allowed to bring the red line within this region.

5. DISCUSSION

In general, in absence of known constraints forward modelling of archaeological magnetic anomalies is a procedure characterized by a large amount of arbitrariety in the choice of a "sufficiently correct" model among the infinite number of magnetization distributions that can generate the observed signal. However, the objective of modelling cannot be attaining a distribution of magnetization that reproduces exactly the geometry of the buried structures. This task would be unachievable even in absence of uncertainty because of the intrinsic ambiguity of potential field data [e.g., Blakely, 1995]. Nevertheless, on the basis of a priori knowledge we can always choose a distribution that is archeologically plausible (or compatible with the expected features), with burial depths that are constrained by the spectral analysis of the magnetic anomaly field [Spector and Grant, 1970], and built with the typical materials of the studied settlement. All the solutions that satisfy these three constraints should be considered equivalent from the geophysical point of view and the selected distribution represents a rigorous prediction of the overall arrangement of the buried structures, which can be falsified by direct excavation.

The techniques described in the previous sections have the primary objective to allow an accurate reconstruction of buried settlements. We showed that error analysis and assessment of the uncertainty are essential steps in the geophysical interpretation (i.e., in terms of physical parameters of buried features) of data that have been acquired using accurate survey techniques. For example, the magnitude of the uncertainty field $\varepsilon(x,y)$ could prevent the interpretation of the anomalies in some areas. In general, the main component of the total uncertainty, ε, of a set of magnetic field observations $T(x,y,t)$ depends from random positioning errors and is proportional to total field gradients. However, there is always a small "background" uncertainty ε_0 that is independent from both field amplitudes and gradients, so that the observed total field uncertainty ε results to be given by:

$$\varepsilon(x,y) = \xi G(x,y) + \varepsilon_0 \tag{7}$$

The secondary component of uncertainty ε_0 includes an uncertainty associated with the instrumental sensitivity, which is generally much smaller than 1 nT depending from the sampling frequency, and the uncertainty related to the statistical fit of the total field data $T(x,y)$ by a polynomial surface $F(x,y)$. The latter quantity is generally very small (less than 3 nT) and can be estimated taking the mean half amplitude of the prediction interval of the regression surface $F(x,y)$.

The total uncertainty can be reduced considerably using specific survey strategies and additional hardware. For example, the positioning errors can be reduced to a negligible value using a motorized total station that follows the sensor [Bruniaux et al., 2018], while a base station for the acquisition of diurnal drift variations can be used to eliminate the component of background uncertainty associated with the variance of the estimated diurnal drift curve. However, some survey strategies are not necessarily appropriate and can introduce systematic errors in the resulting magnetic anomaly field. This could be the case of methods based on the simultaneous acquisition of two total field values by a vertical gradiometer having sensor separation of 1.5 m or greater. In this approach, the total field anomalies are calculated by subtraction of the upper sensor readings, which are considered representative of crustal and core contributions, from the lower sensor values. It is assumed that the vertical pseudogradients provide a good estimate, cleared by diurnal drift variations, of the archaeological magnetic anomalies, because the long–wavelength bandwidth and the external contributions are read simultaneously by the two sensors and have the same value. Therefore, they cancel out in the calculation of the vertical pseudogradient. To test the validity of this approach we simulated the acquisition of pseudogradient data at the Urbs Salvia site by taking the 1.5 m and 2.5 m upward continuations of the total field data used to build the magnetic anomaly grid of Figure 10. Then, we recalculated magnetic anomalies by subtraction of these data from the original data set. A comparison between these anomaly fields and the data in Figure 10 is illustrated in Figure 14 by the two difference maps. The simulated 1.5 m pseudogradient grid has an rms error of 6.93 nT relative to the grid in Figure 10, with a range of [−169.5, 286.9] nT. The simulated 2.5 m pseudogradient has an rms error of 6.11 nT relative to the grid in Figure 10, within the range [−191.4, 294.1] nT. Therefore, this approach does not seem to provide results that can be considered equivalent to the canonical reduction of total field data to magnetic anomalies. Rather, an analysis of these anomalies shows that they strongly accentuate the short wavelengths and can be reproduced by HP filtering of the magnetic anomalies shown in Figure 10. It is also important to note that while magnetic anomalies calculated by Equation (1) have a canonical physical interpretation, because they represent the projection of the anomalous field generated by archaeological features onto the reference field direction [e.g., Blakely, 1995], any other anomaly grid (including residual anomalies calculated by HP filtering) can only provide an approximation of these quantities.

In addition to random errors that contribute to the total uncertainty of magnetic anomaly data, we have mentioned in the previous sections a series of common processing errors that may affect the resulting magnetic maps. For example, errors arising from the selected gridding procedure or from an incorrect choice of the polynomial degree N could affect the physical interpretation of the observed data. In particular, a wrong knitting procedure between the regions that compose the survey area could introduce a serious distortion in the resulting magnetic anomaly map. In
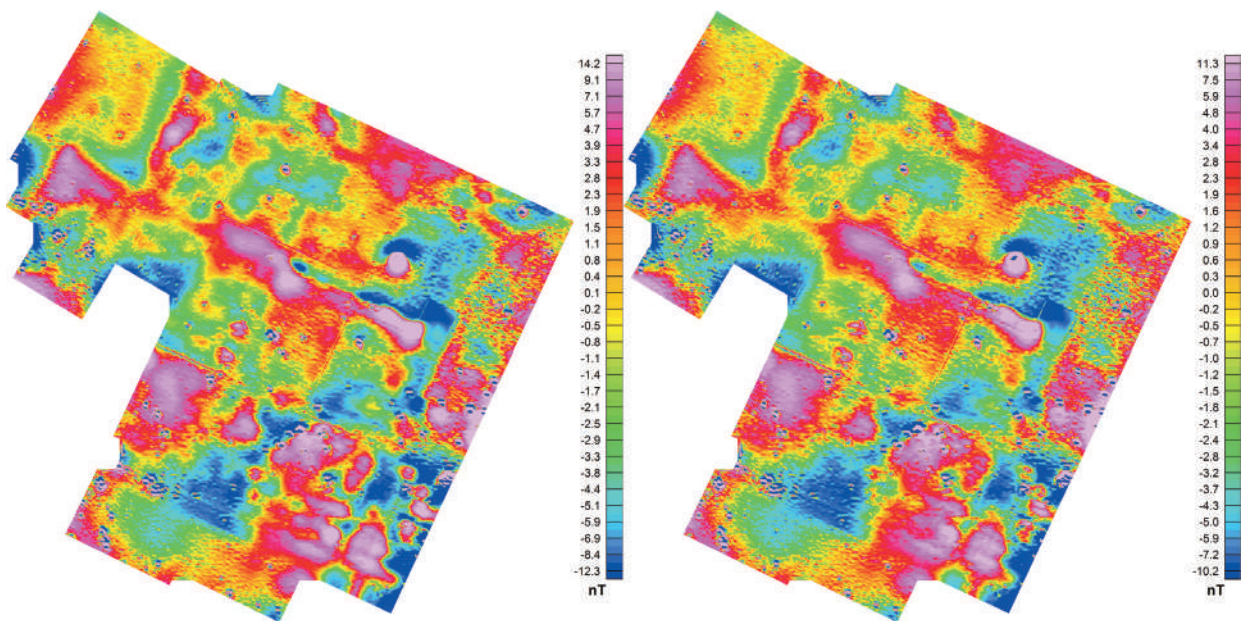

FIGURE 14. Difference grids for the Urbs Salvia Forum area, obtained subtracting magnetic anomaly grids of simulated pseudogradi–ent data with 1.5 m (left) and 2.5 m (right) sensor separation from the magnetic anomaly grid of Fig. 10.

principle, it is possible to combine the data from several regions into a single database and then apply a unique gridding procedure to the whole collection. Although this approach could be appropriate when the objective is the acquisition and analysis of gradient data, in the case of total field surveys, or when total field data must also be extracted from a gradiometer survey, this technique will produce wrong results. In fact, any kind of diurnal correction applied to the data acquired at a single region will remove the time variations of the total field relative to an arbitrary initial level. When at a later time we acquire new data along an adjacent region, the correction will remove time variations of the total field relative to a different initial level, in most cases separated by several tens nT from the reference level of the previous survey. Consequently, a significant gap generally exists along adjacent edges of the regions that compose the survey area. This problem could be partially overcome reducing the total field of the survey regions to magnetic anomalies independently from each other and then applying the knitting procedure directly to the anomaly grids. However, this approach has two flaws. First, even the sophisticated knitting algorithms discussed above could not remove completely the edge effect. Secondly, the best results in the reduction of total field data to magnetic anomalies using the method of Schettino et al. [2018] are obtained performing the analysis at the scale of the whole survey area, not on the individual much smaller regions. Therefore, we believe that the correct approach to the knitting of total field data should start with a procedure of equalization of the survey regions, which shifts each data set except one reference region by a constant value depending on the rms error of the difference grids along overlapped edges. Then, a normal knitting procedure is applied to the equalized total fields of each region. This approach provides the best results when the survey regions overlap along their edges by 1 m.

6. CONCLUSIONS

In the previous sections, we have reviewed common sources of errors and uncertainty in the acquisition and analysis of archaeological magnetic data, considering in particular our own approach to the acquisition and processing of these data, but also discussing other methods. We have also presented a technique for the correct estimation of the errors and the determination of the total uncertainty, which poses a lower bound to the interpretable magnetic anomalies. In our approach, the data are interpreted by a rigorous forward modelling procedure that allows creating a realistic representation of the buried structures. However, in the most favourable condition, non–

uniqueness of forward modelling solutions must be considered to evaluate alternative possibilities. Finally, we have presented an application of our approach to uncertainty assessment, analysis, and modelling of magnetic anomalies through the study of some interesting features of the buried forum area of the Urbs Salvia archaerological site.

REFERENCES

Adisoma, G. S. and M.G. Hester (1996). Grade estimation and its precision in mineral resources: the Jackknife approach. Mining Engineering. 48(2), 84-88.

Aidona, E., R. Scholger, H.J. Mauritsch, and M. Perraki, (2008). Remanence acquisition in a Roman-style gold furnace. Physics and Chemistry of the Earth, Parts A/B/C, 33(6-7), 438-448.

Blakely, R.J. (1995). Potential Theory in Gravity and Magnetic Applications, 441 pp., Cambridge University Press, Cambridge, UK.

Bourges, M., J.L. Mari, and N. Jeannée (2012). A practical review of geostatistical processing applied to geophysical data: methods and applications, Geophys Prospect, 60(3), 400-412.

Bruniaux, G., V. Mathé, F. Lévêque, A. Camus, and V. Ard (2018). Data processing chain for high spatial resolution magnetic survey: Application on the Neolithic site of le Pontet (Charente maritime, France), Archaeol. Prospect., 25(1), 3-16.

Caratori Tontini, F. (2012). Rapid interactive modeling of 3D magnetic anomalies, Computers & Geosciences, 48, 308-315.

Cheesman, S., I.MacLeod and G. Hollyer (1998). A new, rapid, automated grid stitching algorithm, Explor. Geophys, 29(4), 30,1-305.

Cressie, N. (1990). The origins of kriging, Math. Geol, 22(3), 239-252.

Davis, B.M. (1987). Uses and abuses of Cross-validation in geostatistics. Mathematical Geology, 19(3), 241-248.

Desvignes, G., A. Tabbagh and C. Benech (1999). The determination of the depth of magnetic anomaly sources, Archaeol. Prospect., 6(2), 85-105.

Fedi, M., and G. Florio (2003). Decorrugation and removal of directional trends of magnetic fields by the wavelet transform: Application to archaeological

areas, Geophys, Prospect., 51(4), 261-272.

Jacobs, J.A. (1970). Geomagnetic Micropulsations, 179 pp., Springer, Berlin, doi:10.1007/978-3-642-86828-3.

Nabighian, M. N., V.J. S. Grauch, R.O. Hansen., T.R., LaFehr, Y. Li, J.W. Peirce, J. D. Phillips and M.E. Ruder (2005). The historical development of the magnetic method in exploration, Geophysics, 70(6), 33–61.

Parker, R.L. and S.P. Huestis (1974). The inversion of magnetic anomalies in the presence of topography, J. Geophys. Res., 79(11), 1587-1593.

Parker, R.L. (1977). Understanding inverse theory, Annual Review of Earth and Planetary Sciences, 5(1), 35-64.

Powell, A.J., J.G. McDonnell, C.M. Batt, and R.W. Vernon (2002). An assessment of the magnetic response of an iron smelting site. Archaeometry, 44(4), 651-665.

Schettino, A. (2014). Quantitative Plate Tectonics, 403 pp., Springer, Berlin, ISBN 978-3-319-09134-1.

Schettino, A., A. Ghezzi, and P.P. Pierantoni, (2018). Magnetic field modelling and analysis of uncertainty in archaeological geophysics, Archaeol. Prospect., in press, DOI: 10.1002/arp.1729.

Smith, R.S. and M.D. O'Connell (2005). Interpolation and gridding of aliased geophysical data using constrained anisotropic diffusion to enhance trends, Geophysics, 70(5), V121-V127.

Spector, A. and F.S. Grant (1970). Statistical models for interpreting aeromagnetic data, Geophysics, 35(2), 293-302.

Thompson, D.T. (1982). EULDPH: a new technique for making depth estimates from magnetic data computer-assisted, Geophysics, 47, 31–37.

Tomczak, M., (1998). Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW) - cross-validation/jackknife approach, Journal of Geographic Information and Decision Analysis, 2(2), 18-30.

*CORRESPONDING AUTHOR: Annalisa GHEZZI,

University of Camerino, School of Science and Technology,

Geology Division, Camerino (MC), Italy

email: annalisa.ghezzi@unicam.it