FFI-RAPPORT

16/00914

Modelling of civilian ships' ferromagnetic signatures

—
Mads Stormo Nilsson

Modelling of civilian ships' ferromagnetic signatures

Mads Stormo Nilsson

Norwegian Defence Research Establishment (FFI) 11th May 2016

Keywords

Skipssignatur
Regresjonsanalyse
Magnetiske felt
Modellering og simulering

FFI-rapport:
FFI-RAPPORT 16/00914

Project number:
1321

ISBN
P: ISBN 978-82-464-2742-3
E: ISBN 978-82-464-2743-0

Approved by
Morten Nakjem, *Research Manager*
Elling Tveit, *Director*

Summary

In FFI project 1321 *Teknologisk risikoreduksjon for fremtidig minesveip*, the goal is to support the acquisition of future mine countermeasure capability for the Royal Norwegian Navy. A part of this support is recommending performance requirements for future influence sweep systems. We have available a large data set of measured civilian ship signature data from the EDA project SIRAMIS. To determine the necessary performance for new influence sweep systems an analysis of this data set should be performed. The measurement are largely targets of opportunity where ships pass the sensor arrays at different speeds and distances. In order to do an ”apples-to-apples“ comparison we therefore want to produce models of the measured ship signatures, which can then be evaluated under the same conditions. This report details a study evaluating different methods for fitting a model to the magnetic signature data.

The Prolate Spheroidal Harmonic (PSH) model is a mathematical model that describes the magnetic field of a ship as the cumulative effect of a collection of magnetic multipoles placed in a prolate spheroidal coordinate system. In order to get a good description of the magnetic field of a ship using the PSH model, the coefficients determining the strength of each multipole must be found through some fitting procedure.

It is important to avoid overfitting when optimising the model coefficients. The degree of overfitting in a optimised model can be found by evaluating the model’s predictive ability. Many different linear regression methods for optimising the model coefficients have been evaluated based on the predictive ability of the models they produce. The Lasso LARS method was found to give the least amount of overfitting and was chosen to produce models based on the data set. In addition a non-linear method for optimising the size of the prolate spheroidal coordinate system was used.

Magnetic models were produced for the civilian ship signature measurements. Analysis of these ship models will form the basis for future recommendations of influence sweep performance requirements.

Sammendrag

I FFI prosjekt 1321 *Teknologisk risikoreduksjon for fremtidig minesveip* er målet å støtte anskaffelsen av fremtidig minemottiltakskapabilitet for den Kongelige Norske Marinen. En del av denne støtten er anbefaling av krav til ytelse for fremtidige influenssveipesystemer. Vi har tilgjengelig et stort datasett av målte sivile skipssignaturdata fra EDA prosjektet SIRAMIS. For å kunne bestemme den nødvendige ytelsen for nye influenssveipesystemer så bør en analyse av dette datasettet gjennomføres. Målingene er hovedsakelig tatt av tilfeldig forbipasserende skip, hvor de passerer med forskjellige hastigheter og avstander fra sensorene. For å kunne gjøre en riktig sammenligning av signaturdata ønsker vi å produsere modeller av de målte skipsignaturene som kan evalueres under de samme forutsetningene. Denne rapporten beskriver en studie som evaluerer forskjellige metoder for tilpasning av en modell til magnetiske signaturdata.

Den prolatsfæriodiske harmoniske modellen (PSH) er en matematisk modell som beskriver magnetfeltet til et skip som den akkumulerte effekten av en samling magnetiske multipoler plassert i et prolatsfæriodisk koordinatsystem. For å få en god beskrivelse av magnetfeltet til et skip med PSH modellen så må vi finne verdien til koeffisientene som bestemmer styrken på hver multipol gjennom en tilpasningsprosedyre.

Det er viktig å unngå overtilpasning ved optimering av modellkoeffisientene. Graden av overtilpasning i en optimert modell kan finnes ved å evaluere modellens prediksjonsevne. Mange forskjellige lineære regresjonsmetoder for optimering av modellkoeffisientene har blitt evaluert basert på deres evne til å produsere modeller med god prediksjonsevne. Metoden som ga minst overtilpasning var Lasso LARS, som ble valgt til å produsere modeller basert på datasettet. I tillegg ble en ikke-lineær metode for optimering av størrelsen på det prolatsfæriodiske koordinatsystemet brukt.

Magnetiske modeller ble produsert for sivile skipssignaturmålinger. Analyser av disse skipsmodellene vil danne grunnlaget for fremtidig anbefaling av krav til ytelse av influenssveip.

Contents

1 Introduction

In FFI project 1321, "*Teknologisk risikoreduksjon for fremtidig minesveip*", the goal is to support the acquisition of future mine countermeasure capability for the Royal Norwegian Navy. New magnetic and acoustic influence sweep systems will be a part of this capability.

Before acquiring a sweep system it is important to specify the performance requirements of the system. Since high requirements often translate into high costs, we want to find the minimum requirements for the systems that still produces satisfactory performance.

Clearing naval mines to allow for transit of civilian vessels through a sea zone is one of the primary requirements for a sweep system. In order to find the necessary performance for this task we must look at the signatures of civilian vessels and compare with the signatures produced by a sweep system. In this work we will focus on analysis and modelling of the magnetic signatures of civilian vessels.

Signature **R**esponse **A**nalysis on **M**ulti-**I**nfluence **S**ensors[1] (SIRAMIS) was a co-operative EDA project between Norway, Sweden, France, Germany, Netherlands, Poland and Spain. The main goal of the SIRAMIS project was to increase knowledge about ship signatures and how multi influence mines react to the near field. The first step in this work was the collection of signature measurements of civilian shipping. Each nation conducted measurement campaigns with its own equipment and/or measurement ranges. The measured signatures have been collected in a database and distributed to the member nations.

1.1 Modelling

In this analysis we have used signature data measured and collected through the SIRAMIS project [1]. The Prolate Spheroidal Harmonic model has been chosen to model the measured magnetic ship signature data. A detailed analysis of different model fitting methods has been done and magnetic ship models have been generated with the chosen method. These ship models have been analyzed to find scaling factors and correlations between the magnetic signature and metadata such as ship types, size, tonnage, etc. The models have been compared with sweep models to find the necessary sweep performance for the vessels.

In order to do an "apples to apples" comparison of different ships we need to know the ships' magnetic signatures at the same depth and distance from the sensor. We can achieve this by constructing models, which are used to estimate the ship's magnetic field at identical depth and distance from the hull.

It is possible to model a ship's magnetic field by creating a smaller scale physical model of its hull. This approach requires in-depth knowledge about the materials and geometry of the hull, which is not feasible when modelling a large collection of merchant vessels. We have decided to use a semi-empirical model where the coefficients of a mathematical model are fitted to the measured magnetic signatures, and are then used to extrapolate the signatures to a standard depth and distance.

[1] EDA project A-919-ESM1-GP.07, FFI project 1224

While this approach allows us to generate models for a large number of vessels in a short time, it also introduces the possibility of overfitting the model. By overfitting we mean that the model coefficients are too fine tuned, making the model describe not only the ship's magnetic field, but also any noise or measurement error. An overfit model typically describes the measured part of the field well, but gives a very bad description of the field outside the measurement points. Care must be taken in choice of fitting procedure to reduce overfitting, and preserving the models' estimating power.

In the next chapters we will describe the mathematical model chosen to describe the magnetic ship signatures. We will also describe different strategies for choosing correct model parameters in order to get a good description, and how we validate the model after the parameters have been chosen.

2 Prolate spheroidal harmonic model

The scalar potential Φ of a magnetic field must satisfy Laplace's equation [2],

$$\nabla^2\Phi = 0. \tag{2.1}$$

By solving Laplace's equation in spherical coordinates we get the general solution for Φ outside a spherical shell

$$\Phi(r,\theta,\varphi) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\left[c_{nm}\cos(m\varphi) + s_{nm}\sin(m\varphi)\right] P_n^m(\cos\theta)/r^{n+1}, \tag{2.2}$$

where P_n^m is the associated Legendre function of the first kind. c_{nm} and s_{nm} are the coefficients for the harmonic functions and can be interpreted as weight factors of magnetic multipoles with degree n and order m.

We can obtain the magnetic field from the scalar potential

$$\vec{B} = -\mu_0\nabla\Phi. \tag{2.3}$$

By choosing the correct values for the coefficients, this model can be used to describe the magnetic field of a ship. It is important however to keep in mind the limitations of the model. In order to get a proper representation of the near field we want the shell to enclose the entire ship. This means that for a ship with length l we can only generate models using data measured with sensors placed at a distance from the ship $r > l/2$. Furthermore, once generated the model can only be meaningfully evaluated for $r > l/2$. Since a ship's length is typically much larger than its beam or height, a spherical coordinate system is not necessarily the best choice when modelling its magnetic field.

The typical form of a ship can be approximated as an ellipsoid with length l and height h. A prolate spheroidal (PS) coordinate system produces ellipsoidal surfaces and will with its "cigar" shape be a potentially better choice for modelling the ship. If we place the ship inside an ellipsoid with length $a \approx l$ and height $b \approx h$, the magnetic field can be reconstructed outside this PS surface [3]. The case where the magnetic source is placed inside either a spherical or a spheroidal surface is shown in Figure 2.1. With a PS coordinate system one can describe the magnetic field closer to the ship than for a equivalent spherical coordinate system.

A PS coordinate system is defined by a focal length f, and has the coordinates ξ, η, φ as shown in Figure 2.2. The coordinate limits are $1 \le \xi < \infty$, $-1 \le \eta \le 1$ og $0 \le \varphi < 2\pi$. If we hold ξ constant we get ellipsoidal surfaces, constant η giver hyperbolic surfaces and constant φ gives planes along the given angle.

The following transformations takes us from the Cartesian to the PS coordinate system,

$$\begin{aligned}\xi &= \frac{r_2 + r_1}{2f}\\ \eta &= \frac{r_2 - r_1}{2f}\\ \varphi &= \arctan\frac{y}{z},\end{aligned} \tag{2.4}$$

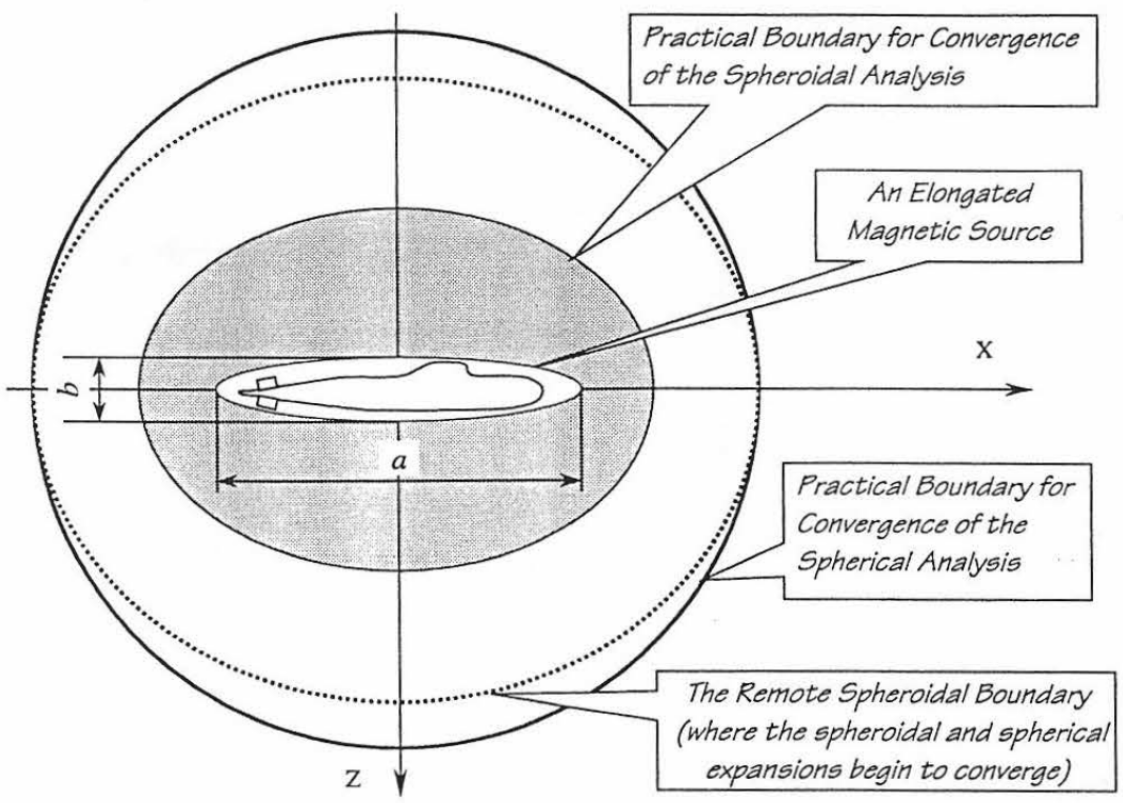


Figure 2.1 *The limits for modelling the ship signature in a spherical and prolate spheroidal coordinate system. Figure taken from [3]*

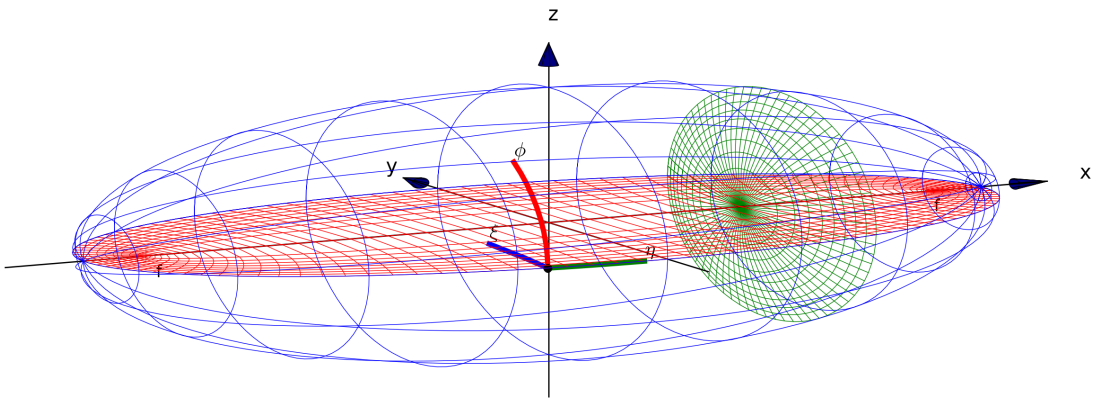


Figure 2.2 *Prolate spheroidal coordinate system. Holding ξ, ϕ or η constant gives the surfaces in blue, red or green respectively*

where

$$r_1 = \sqrt{(x-f)^2 + y^2 + z^2}$$
$$r_2 = \sqrt{(x+f)^2 + y^2 + z^2}. \tag{2.5}$$

If we solve Laplace's equation in the PS coordinate system we get the solution [4]

$$\Phi(\xi, \eta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \left[c_{nm} \cos(m\varphi) + s_{nm} \sin(m\varphi) \right] P_n^m(\eta) Q_n^m(\xi), \tag{2.6}$$

where P_n^m and Q_n^m are the associated Legendre functions of first and second kind respectively, with degree n and order m. c_{nm} and s_{nm} are the model coefficients, and can as in the spherical case be interpreted as weight factors of magnetic multipoles with degree n and order m.

Using (2.3) and (2.6) we get the magnetic field from the **P**rolate **S**pheroidal **H**armonic (PSH) model. By putting one of the coefficients c_{nm} and s_{nm} to one and the rest to zero, we get the magnetic field generated by the multipole associated with that coefficient. The generated field from a PSH model can be seen as a superposition of the different fields associated with the coefficients c_{nm} and s_{nm}.

In Table 2.1 and 2.2 we see the total magnetic fields $B_T = \sqrt{\vec{B}_x^2 + \vec{B}_y^2 + \vec{B}_z^2}$ generated by the multipoles associated with the different PSH coefficients. We can see the dipole moments for $n = 1$, the quadropole moments for $n = 2$ and the octopole moments for $n = 3$. For increasing $n - m$ we get an increasing number of oscillations in $\vec{B}_x, \vec{B}_y, \vec{B}_z$ in the longitudinal ship direction, and for $n = m$ we get more oscillation in the athwartship direction.

Table 2.1 *Magnetic fields generated by multipoles associated with the c_{nm} PSH coefficients for $1 \leq n \leq 3$, $0 \leq m \leq 3$. Field evaluated in the xy-plane at 30 m depth with a PS focal length of 150 m.*

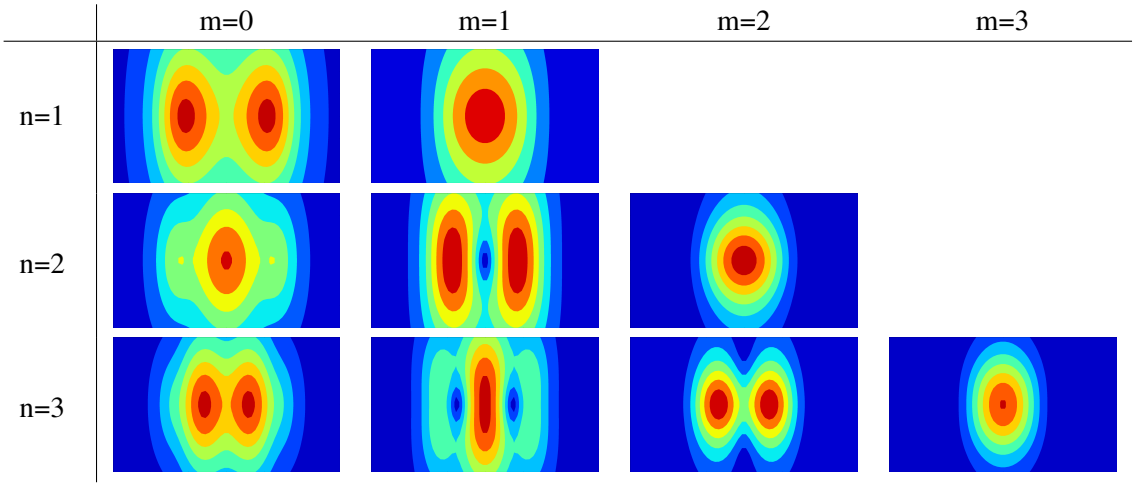
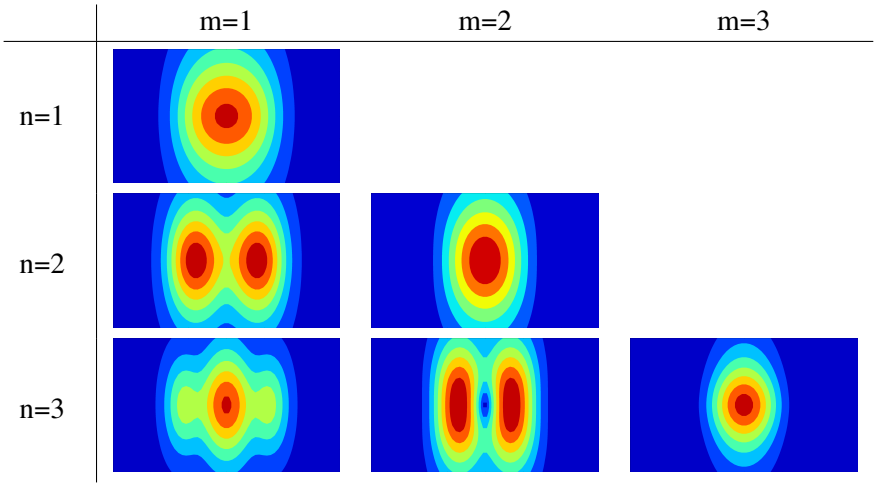


Table 2.2 *Magnetic fields generated by multipoles associated with the s_{nm} PSH coefficients for $1 \leq n \leq 3$, $1 \leq m \leq 3$. Field evaluated in the xy-plane at 30 m depth with a PS focal length of 150 m.*



3 Model fitting

In order to get a good description of the magnetic field of a target ship the choice of PSH model coefficients is essential. The process of choosing coefficients so that the model describes the ship's magnetic field is called inverse modelling, and is done through a numerical fitting routine. We use k coefficients to fit a PSH model to a measurement with l measurement points. For a given center point and focal length for a PS coordinate system we can find the coefficients with a linear regression method [5].

3.1 Linear regression methods

We put the coefficients from the PSH model in a vector $\vec{\beta}$, mapping the coefficients β_i to c_{nm} and s_{nm} such that $\beta_0 = c_{00}$, $\beta_1 = c_{10}$, $\beta_2 = c_{11}$, $\beta_3 = s_{11}$, and so on

$$\vec{\beta} = [\beta_1, \beta_2, ..., \beta_k]^T = [\underbrace{c_{00}, c_{10}, c_{11}, s_{11}, c_{20}, c_{21}, s_{21}, ...}_{k \text{ elements}}]^T. \tag{3.1}$$

We wish to find $\vec{\beta}$ such that it minimises the total square error

$$S(\vec{\beta}) = \left\| \vec{y} - \vec{\mu}(\vec{\beta}) \right\|^2 = \sum_{i=1}^{l} \left(y_i - \mu_i(\vec{\beta}) \right)^2, \tag{3.2}$$

where $\vec{y} = [y_1, y_2, ..., y_l]$ are the measured values of the x-, y- and z-components of the ship's magnetic field in the measurement points. $\vec{\mu}(\vec{\beta}) = [\mu_1(\vec{\beta}), \mu_2(\vec{\beta}), ..., \mu_l(\vec{\beta})]$ are the model predictions at the same measurement points given as

$$\vec{\mu}(\vec{\beta}) = X\vec{\beta}, \tag{3.3}$$

where X is a $l \times k$ matrix, with elements defined as

$$X_{ij} = \frac{\partial B(\xi_i, \eta_i, \varphi_i)}{\partial \beta_j}. \tag{3.4}$$

We can split X into a set of vectors $X = [\vec{x}_1, \vec{x}_2, ..., \vec{x}_k]$, where the vectors are given as $\vec{x}_j = [X_{1j}, X_{2j}, ..., X_{kj}]^T$. $\vec{x}_j$ are the basis vectors of the parameter space spanned by the PSH multipoles.

A number of methods can be used to optimise the coefficients $\vec{\beta}$ so that the resulting model describes the measured ship magnetic field. We will present and evaluate some of the methods. Our main focus will be how the methods deal with the problem of overfitting. In general a model using more parameters will get a better fit to the training data, but will results in greater overfitting and degradation of the ability to generalise.

3.1.1 Ordinary least squares

The **O**rdinary **L**east **S**quares [5] (OLS) fit method simply minimises the total square error. The solution coefficients are given by

$$\vec{\beta} = (X^T X)^{-1} X^T \vec{y}. \tag{3.5}$$

This method gives the best minimisation of (3.2), but will also be susceptible to overfitting, especially when using many coefficients. We use OLS as the baseline against which we compare the other methods.

3.1.2 Truncated singular value decomposition

We can introduce truncated **S**ingular **V**alue **D**ecomposition [6] (SVD) of the basis matrix into the least square fit in order to reduce the effect of noise in the data set. X is decomposed into

$$X = USV^T, \tag{3.6}$$

where S is an $l \times k$ diagonal matrix whose values s_{ii} are the squares of the eigenvalues of both XX^T and $X^T X$. U is a $l \times l$ matrix whose columns $\vec{u}_i$ are the eigenvectors of XX^T. V is a $k \times k$ matrix whose columns $\vec{v}_i$ are the eigenvectors of $X^T X$. The values s_{ii} are called the singular values of X. If we insert (3.6) into (3.5) we get

$$\vec{\beta} = VS^{-1}U^T y = \sum_i \frac{\vec{u}_i \cdot \vec{y}}{s_{ii}} \vec{v}_i. \tag{3.7}$$

We write $\vec{y} = \vec{y}' + \vec{e}$, where $\vec{e}$ is the effect of noise and measurement error, and $\vec{y}'$ is the real magnetic field we want to model, so that we get

$$\vec{\beta} = \sum_i \frac{\vec{u}_i \cdot (\vec{y}' + \vec{e})}{s_{ii}} \vec{v}_i. \tag{3.8}$$

The error from noise in the ordinary least square solution is

$$\vec{\beta} - \vec{\beta}' = \sum_i \frac{\vec{u}_i \cdot \vec{e}}{s_{ii}} \vec{v}_i, \tag{3.9}$$

where $\vec{\beta}'$ is the OLS solution for $\vec{y}'$. We can see that for small singular values s_{ii} the error will be greatly magnified. In general small singular values can be seen as a sign of a poorly conditioned basis matrix. We can attempt to reduce the error by setting singular values that are less than c to zero (truncating) in (3.6), where c is chosen so that only significant values of s_{ii} are kept. Significant values is taken to mean the s_{ii} that are needed to describe the field $\vec{y}'$. If there are no singular values smaller than c the truncated SVD solution becomes the ordinary least squares solution.

3.1.3 Regularisation

Using a regularisation procedure is a common method for avoiding overfitting in a regression analysis. A regularisation procedure introduces additional information into the fit in an attempt to

solve an ill-posed problem. We limit the size of the coefficients in the fit by minimising (3.2) with the constraint

$$R(\vec{\beta}) \leq t, \tag{3.10}$$

where $R(\vec{\beta})$ is a regularisation function and t is a tuning parameter. If we introduce a Lagrange multiplier α [5], the expression to be minimised becomes

$$S'(\vec{\beta}) = \left\|\vec{y} - \vec{\mu}(\vec{\beta})\right\|^2 + \alpha R(\vec{\beta}), \tag{3.11}$$

where α is a regularisation tuning parameter with inverse value of t. For a given fitting problem and regularisation function, we have a continuum of possible solutions determined by the value of α in the range $[0, \infty)$. $\alpha = 0$ gives the ordinary least squares solution, while $\alpha \to \infty$ gives $\vec{\beta}$ such that $R(\vec{\beta}) \to 0$.

3.1.3.1 Ridge

Ridge [7] is a regularisation method where the squared L_2-norm is used as the regularisation function

$$R(\vec{\beta}) = ||\vec{\beta}||_2^2 = \sum_{i=1}^{k} \beta_i^2. \tag{3.12}$$

From (3.11) we see that $\alpha \to \infty$ gives $\vec{\beta} \to \vec{0}$. The choice of α in a Ridge fit will therefore be a choice between small model coefficients and a linear fit. A set of coefficients where some β_i are large and some are small will be penalised more heavily than a set of coefficients where all β_i are of medium size. This means that the Ridge method might discard some potential fit results where the values of β_i varies greatly with i.

3.1.3.2 Lasso

The **L**east **a**bsolute **s**hrinkage and **s**election **o**perator (Lasso) method [8] is a regularisation method where the L_1-norm is used as the regularisation function

$$R(\vec{\beta}) = ||\vec{\beta}||_1 = \sum_{i}^{k} |\beta_i| \,. \tag{3.13}$$

The Lasso method has many of the same properties as the Ridge method. The choice of α is a choice between small coefficients and a linear fit, with $\alpha \to \infty$ giving $\vec{\beta} \to \vec{0}$. Unlike Ridge it will not discard solutions just because there is a large variation in the size of β_i, since Lasso penalises the solution only for the absolute size of the coefficients.

An important property of the Lasso method is that coefficients that do not contribute with explanatory power to the model are put to exactly zero. This is in contrast to ordinary least squares and other regularisation methods where coefficients that do not contribute usually will be set to a small non-zero value. We can see why if we look at a simple example with two coefficients.

In Figure 3.1 we can see in gray the constraint imposed on the fit by (3.10) for Lasso (left) and Ridge (right). The point $\vec{\beta}$ marked is the solution for an ordinary least squares method, and

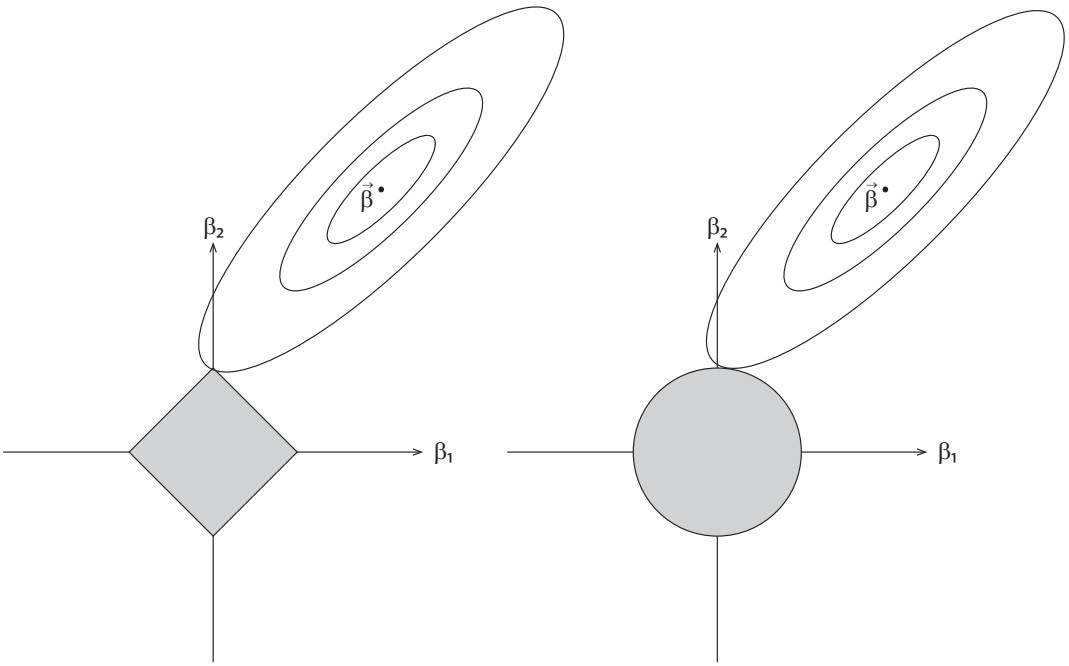


Figure 3.1 *Comparison of Lasso (left) and Ridge (right) regularisation. Figure reproduced from [8].*

points with constant error are shown along the contours. The regularisation solution is found in the intersect between the contour with smallest possible error and the gray area defined by $R(\vec{\beta}) \leq t$. In our example Lasso gives $\beta_1 = 0$, while Ridge gives $\beta_1 \neq 0$. In general there is a larger probability that the Lasso gives a solution where one or more of the coefficients β_i is put to zero, while the corresponding solution for Ridge will give a small non-zero value. Generally more and more β_i becomes zero with increasing α.

3.1.4 LARS

Least **A**ngle **R**egression (LARS) [9] is a method for building up the coefficient vector $\vec{\beta}$ one element at a time in an iterative process. After j iterations of the algorithm we will have found j non-zero elements of $\vec{\beta}$. We define $\vec{\mu}_{L_j}$ as the model prediction after j iterations, and the correlations of the basis vectors with the unmodeled parts of the measurements as

$$c_i(\vec{\mu}_{L_j}) = \vec{x}_i \cdot (\vec{y} - \vec{\mu}_{L_j}). \tag{3.14}$$

We start the process with the initial prediction $\vec{\mu}_{L_0} = \vec{0}$ and calculate the correlations c_i between the basis vectors $\vec{x}_i$ and the measurement $\vec{y}$

$$c_i(\vec{\mu}_{L_0}) = \vec{x}_i \cdot \vec{y}. \tag{3.15}$$

We order the basis vectors according to their correlation, calling the most correlated $\vec{x}_{L_1}$, the second most correlated $\vec{x}_{L_2}$ and so on. The first step of the algorithm is to add $\vec{x}_{L_1}$ to the solution

$$\vec{\mu}_{L_1} = \vec{\mu}_{L_0} + \gamma_1 \vec{x}_{L_1}, \tag{3.16}$$

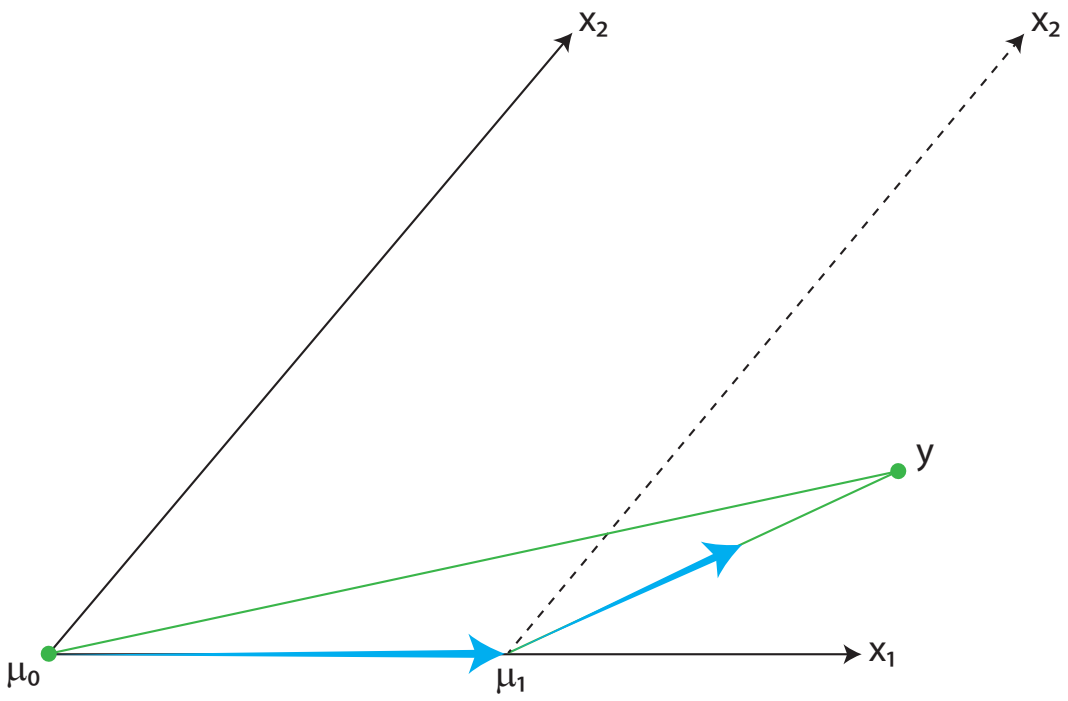


Figure 3.2 *Geometric interpretation of the LARS method. Figure reproduced from [9]*

where vi choose γ_1 such that $\vec{x}_{L_1}$ and $\vec{x}_{L_2}$ are equally correlated with $\vec{y} - \vec{\mu}_{L_1}$, meaning

$$c_{L_1}(\vec{\mu}_{L_1}) = c_{L_2}(\vec{\mu}_{L_1}) = \vec{x}_{L_1} \cdot (\vec{y} - \vec{\mu}_{L_1}) = \vec{x}_{L_2} \cdot (\vec{y} - \vec{\mu}_{L_1}). \tag{3.17}$$

We can ensure that this keeps beings the case in the next step by moving in direction of $\vec{x}_{L_1} + \vec{x}_{L_2}$

$$\vec{\mu}_{L_2} = \vec{\mu}_{L_1} + \gamma_2 \frac{\vec{x}_{L_1} + \vec{x}_{L_2}}{||\vec{x}_{L_1} + \vec{x}_{L_2}||}. \tag{3.18}$$

We generalise the process by adding basis vectors step by step with the formula

$$\vec{\mu}_{L_i} = \vec{\mu}_{L_{i-1}} + \gamma_i \frac{\sum_{j=1}^{i} \vec{x}_{L_j}}{\left|\left|\sum_{j=1}^{i} \vec{x}_{L_j}\right|\right|}. \tag{3.19}$$

In every step γ_i is choosen such that $\vec{x}_{Li}$ is equally correlated with $\vec{y} - \vec{\mu}_{Li}$ as all $\vec{x}_{L_{j<i}}$.

We can visualise the method geometrically in Figure 3.2, where we look at the case with two basis vectors $\vec{x}_1$ and $\vec{x}_2$. We see that the angle between $\vec{x}_1$ and $\vec{y}$ is smaller than the angle between $\vec{x}_2$ and $\vec{y}$. The name least angle regression comes from the fact that the most correlated basis vector is the vector with the smallest angle to $\vec{y}$. The LARS solution is to choose $\vec{\mu}_1$ such that the angle between $\vec{y} - \vec{\mu}_1$ and $\vec{x}_1$ is equal to the angle between $\vec{y} - \vec{\mu}_1$ and $\vec{x}_2$. $\vec{\mu}_2$ is chosen such that $\vec{\mu} = \vec{\mu}_1 + \vec{\mu}_2 = \vec{y}$.

3.1.5 Lasso LARS

Lasso and LARS can be combined into a single algorithm. If we for every step of the LARS algorithm calculate the L_1 norm of the current $\vec{\beta}$ we see that we are in fact generating a whole set of solutions with different values for $||\vec{\beta}||_1$. Similarly we can do multiple calculations with the Lasso method for different α to calculate solutions with different values for $||\vec{\beta}||_1$. If we compare the Lasso and LARS solutions with equal $||\vec{\beta}||_1$ we will find that most of the time they will be equal, and surprisingly if we make a small modification to the LARS method they will always be equal.

It can be shown [9] that if we for each iteration of the LARS method demand $\text{sign}(\beta_{L_i}) = \text{sign}(c_{L_i})$, the solution for any given step will be a Lasso solution for a given value of α. If for any step $\text{sign}(\beta_{L_i}) \neq \text{sign}(c_{L_i})$ the chosen basis vector is not considered for that iteration and the second most correlated basis vector is used instead.

3.2 Akaike's information criterion

The Akaike information criterion (AIC) is a measure of the relative quality of a statistical model [10]. The AIC value of a model is given as:

$$AIC = 2k - 2\ln(L) \tag{3.20}$$

where k is the number of coefficients in the model, and L is the maximum of the likelihood function of the model. In the special case of least squares we have [11]

$$AIC = 2k + l\ln(\sigma^2), \tag{3.21}$$

where $\sigma^2 = S(\vec{\beta})/l$, and l is the number of measurements points.

The AIC value measures the goodness of a fit penalised by the number of coefficients and can be used as a model selection tool. A modified Ridge, Lasso or Lasso LARS method can use the AIC value to set the fitting parameter α.

3.3 Selecting the number of coefficients

Up to now we have talked about using different fitting methods to estimate values for k coefficients. For a real problem we also have to find the value of k. If we limit n and m to maximum values of $n_{\max}$ and $m_{\max}$ respectively, (2.6) becomes

$$\Phi(\xi,\eta,\varphi) = \sum_{n=0}^{n_{\max}} \sum_{m=0}^{m_{\max}} \left[c_{nm}\cos(m\varphi) + s_{nm}\sin(m\varphi)\right] P_n^m(\eta) Q_n^m(\xi). \tag{3.22}$$

The number of oscillations in a magnetic moment over the η coordinate in the scalar potential is

$$\nu = n - m + 1 - \delta_m, \tag{3.23}$$

where δ_m is the Kronecker delta

$$\delta_m = \begin{cases} 0, & \text{if } m \neq 0 \\ 1, & \text{if } m = 0. \end{cases} \tag{3.24}$$

The number of measurement points we have available limits the maximum values of m and ν [12]. The values supported by the measurement data are

$$\begin{aligned} m_{\max} &\leq N_y \\ \nu_{\max} &\leq N_x, \end{aligned} \tag{3.25}$$

where $\nu_{\max}$ is the maximum value of ν, N_x is the number of measurement points in the longitudinal ship direction within one ship length, and N_y is the number in the athwartship direction.

For a given m, $\nu_{\max}$ will define $n_{\max}$. Finding $m_{\max}$ and $\nu_{\max}$ will unambiguously define the size of $\vec{\beta}$. The fit is performed for all combinations of $m_{\max}$ in $[0, N_y]$ and $\nu_{\max}$ in $[1, N_x]$, and for each combination the AIC value is calculated. The $m_{\max}$, $\nu_{\max}$ combination that gives the best AIC value is chosen for the fit.

3.4 Coordinate system optimisation

So far we have looked at the fit with the assumption that we have a given focal length f and center point c of the coordinate system. In reality a good fit is dependent on f and c being chosen based on the geometry of the ship. Since we require a defined coordinate system to set up the basis matrix X, the parameters f and c cannot be included in the linear fit.

We optimise the geometry of the coordinate system using a non-linear least squares method with f and c as its only two parameters. The evaluation function for the non-linear fit generates the basis matrix X for the given parameters and runs a linear fit for that matrix, with the error being returned to the non-linear routine as the error for the chosen f and c. In other words, we choose the focal length and center point that gives us the best linear fit of the PSH coefficients.

With some linear methods we see a tendency for overfitting by choosing very large values for the focal length. We therefore try an alternative regularisation method where we punish the non-linear fit doubly for the error in the dipole coefficients

$$S(\vec{\beta}) = ||\vec{y} - \vec{\mu}||^2 + ||\vec{y} - \vec{\mu}_{\text{dipole}}||^2, \tag{3.26}$$

where

$$\vec{\mu}_{\text{dipole}} = \sum_{i=1}^{3} \vec{x}_i \beta_i. \tag{3.27}$$

The magnetic field generated by the model will decay faster with distance for higher order multipoles, and the far-field is therefore dominated by the dipole contribution from the three first coefficients in (2.6). Double error from the dipole contribution is therefore chosen to nudge the method in the direction of giving the best possible description of the far-field. In the following analysis both methods have been used, being called non-linear least squares with or without Double Dipole Error (DDE).

4 Cross-validation

The measure of a good model is not how well it describes the training data $\vec{y}$ but its predictive power. Predictive power is the measure of how well the model describes data that was not used for training. To assess the predictive power of a model we compare the model prediction to a subset of the data which was not used to train it. This is performed several times, with different subsets of the data used for testing, a process called cross-validation. If a model has a good fit error but a bad predictive power it is a sign that it has been overfit. The goal of the cross-validation analysis is to find the fitting method that generates models with the best possible predictive power.

In the measurements from the Herdla measurement range we typically have many passages and measurements from sensors on four different depths. This gives us more information about the signature than we have for the other SIRAMIS measurements. We therefore use the Herdla measurements to do a cross-validation analysis of the fitting methods.

The root mean square error (RMSE) is the measure of how well the model describes the measurements

$$\text{RMSE} = \sqrt{\frac{S(\vec{\beta})}{\text{DoF}}} = \sqrt{\frac{\sum_{i=1}^{l} r_i^2}{\text{DoF}}}, \tag{4.1}$$

where $r_i = y_i - \mu_i$ is the residual for each measurement point and DoF are the degrees of freedom, defined as the number of measurement points l minus the number of non-zero coefficients in the model.

4.1 Method comparison

Three ships measured at Herdla were chosen and fits of the magnetic ship signatures to PSH models were performed with different fitting methods.

We define the total error over all the vessels for a method as

$$\text{Total error} = \sqrt{\frac{\sum_i^n \text{RMSE}_i^2}{n}}, \tag{4.2}$$

where RMSE_i is the error in the fit for a given vessel and n is the number of vessels. Similarly the total error for the cross-validation data has also been calculated. The results from the cross-validation analysis are given in Table 4.1.

Ordinary least squares has the smallest total fitting error, as expected as OLS ensures mathematically that we get the best possible fit to the training data. However, the good fit comes at the cost of overfitting the model as indicated by the relatively large cross-validation error. The degree to which the other methods balance the goodness of the fit versus model complexity can be seen by comparing their results to the OLS results. If we use Double Dipole Error in the non-linear fit it greatly reduces the total cross-validation error, while slightly increasing the total fitting error. This shows that using DDE might be a viable regularisation method on its own.

Table 4.1 *Cross-validation of least squares fitting methods. LS=Least Squares, DDE=Double Dipole Error*

Non-linear method	Linear regression method	Total fitting error	Total cross-validation error
LS + DDE	Lasso LARS	20.81	87.66
LS	Lasso LARS	20.70	87.93
LS + DDE	Ridge	18.57	107.65
LS + DDE	Ordinary least squares	18.43	149.09
LS + DDE	Truncated SVD	18.43	149.09
LS	Ridge	17.42	692.31
LS	Truncated SVD	17.34	770.38
LS	Ordinary least squares	17.34	770.42

Truncated SVD performs no better than OLS, having a similar cross-validation error. One interpretation of this result is that the errors removed by truncating any singular values were not significantly affecting the fit. The improvement from using DDE is similar to OLS.

Ridge regularisation has a slightly larger total fitting error than ordinary least squares but still has a large total cross-validation error. If we use DDE in the non-linear fit the total cross-validation error decreases greatly.

We see that the Lasso LARS method has the largest total fitting error and the smallest total cross-validation error of all the methods. It is also the only method that does not benefit greatly from introducing DDE into the non-linear fit. This can be seen as an indication that the Lasso regularisation by itself prevents the overfitting.

The reason Ridge performs so much worse than Lasso LARS must be sought in the differences between the methods, mainly the choice of whether to use the L_1- or L_2-norm of the coefficients to regularise the fit. The Lasso tends to drive non-contributing coefficients to exactly zero, while Ridge sets them to almost zero. A coefficient set to almost zero may have a negligible contribution to the magnetic field in the training data but may contribute significantly outside, which might explain why Lasso LARS achieves a much smaller cross validation error than Ridge. Another contributing factor could also be the L_2-norm's tendency to favour coefficients of similar size, which could cause problems if the magnetic field produced by one multipole moment in reality describes most of the data. In Figure 4.1 and Figure 4.2 the distribution of coefficient values in Ridge and Lasso LARS can be seen. We can see that Ridge tends to use more nonzero coefficients than Lasso LARS, and of those coefficients many are close to zero in value.

Double Dipole Error generally gives a better cross-validation error for the linear methods and significantly improves the methods that do not have any other regularisation. Lasso LARS does not benefit significantly from using DDE.

Based on the results of the cross-validation analysis we choose to use Lasso LARS as the fitting method in the rest of the analysis. Using DDE does not give a large improvement for Lasso LARS, we therefore choose to use normal least squares method for the non-linear fit.

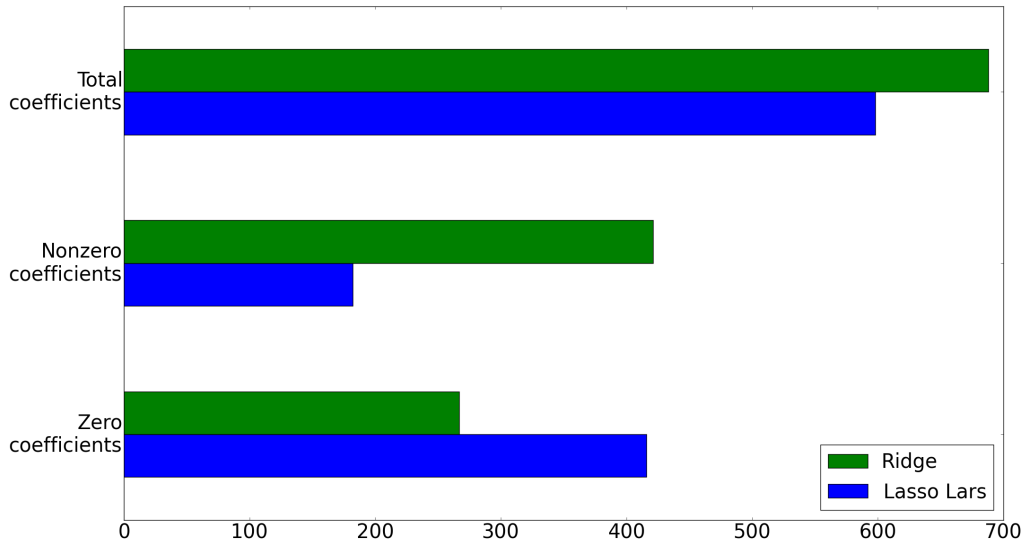


Figure 4.1 *Coefficients used by the Ridge and Lasso LARS methods*

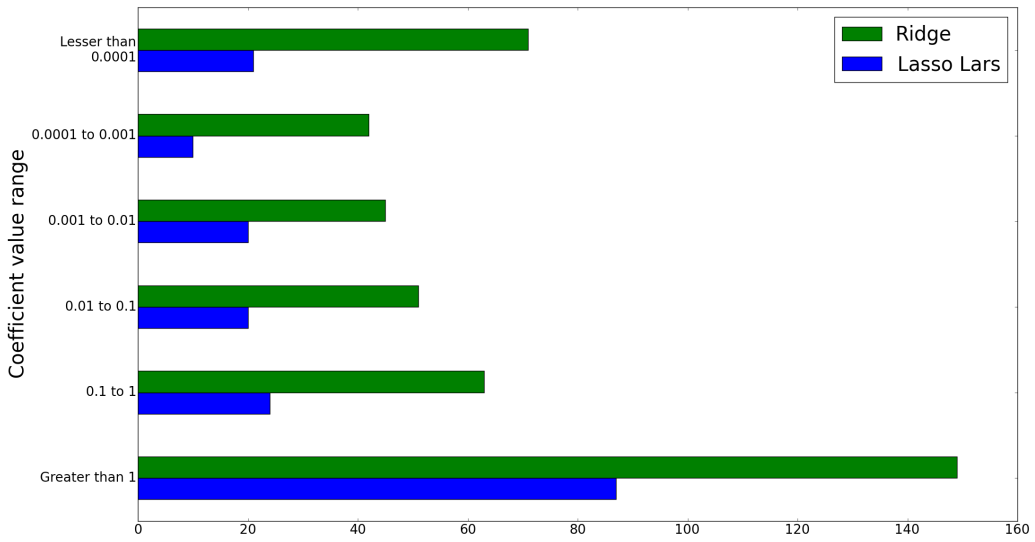


Figure 4.2 *Value distribution of non-zero coefficients used by the Ridge and Lasso LARS methods*

4.2 Tuning Lasso LARS

We know from Section 3.1.5 that each iteration of the Lasso LARS method corresponds to a Lasso fit with a given regularisation parameter α. As the number of iterations increases the value of α decreases. Potentially we can end up with $\alpha \approx 0$ when the fit is completed, and we have an ordinary least squares solution. For the numerical implementation we choose an $\alpha_{\min}$ which is the smallest value of α the method will try. If we after j iterations have $\alpha_j < \alpha_{\min}$ the method ends at that iteration. An additional ending criterion for the method is if we for a given iteration get a larger correlation than for an earlier iteration $c_{Li} < c_{L(i+1)}$. If this happens it means that the numerical error is similar in size to the remaining correlations and we can't improve the model further.

We have performed a cross-validation analysis on the Lasso LARS method in order to investigate its dependency on $\alpha_{\min}$. We see in Figure 4.3 that the total fit error and the total cross-validation error varies little in the range $10^{-6} < \alpha_{\min} < 10^{-2}$. With $\alpha_{\min} > 10^{-2}$ we get increasing fit and cross-validation errors with stabilization for $\alpha_{\min} > 10^{1}$. The explanation can be seen in the graphs showing the maximum and minimum number of non-zero coefficients used in models fit with a given $\alpha_{\min}$. The number of non-zero coefficients decrease with increasing $\alpha_{\min}$, and for $\alpha_{\min} \geq 10^{0}$ some of the models have all coefficients set to zero. In the range $\alpha_{\min} \geq 10^{1}$ all models have all coefficients set to zero. In the case of $\alpha_{\min} = 10^{0}$ we have a larger cross-validation error than for $\alpha_{\min} > 10^{0}$, showing that too few coefficients can give a larger cross-validation error than a model with all coefficients zero. As we want a model that is as simple as possible and with the largest possible predictive power, $\alpha_{\min} = 10^{-2}$ is chosen when fitting models to our data set.

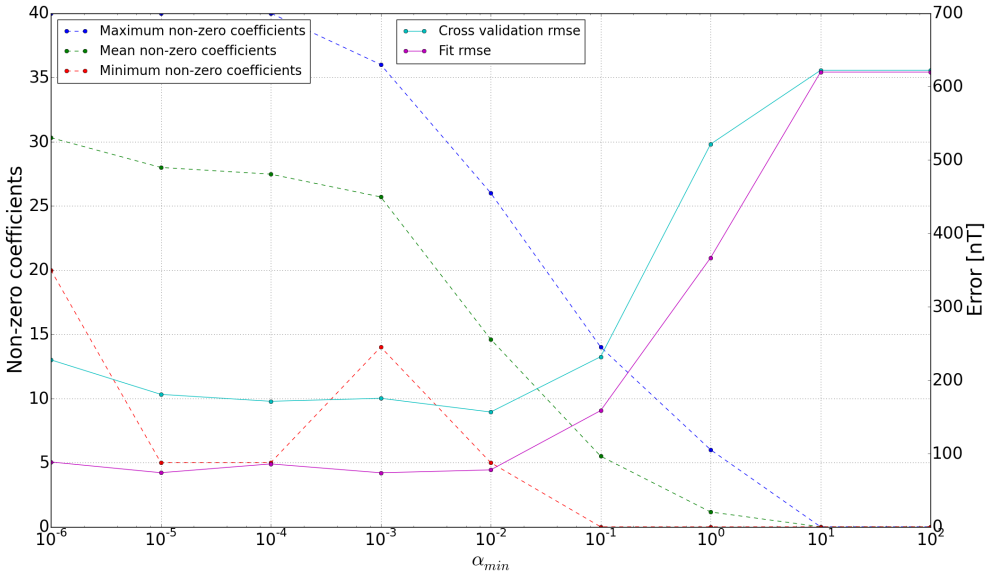


Figure 4.3 *The dependency of Lasso LARS on the regularisation parameter α_{min}.*

4.3 Using the Akaike information criterion to select the Lasso fitting parameter

As an alternative to tuning Lasso LARS manually we can choose the α that minimises the AIC value described in Section 3.2. The cross validation results for this approach and the regular Lasso LARS is compared in Table 4.2. We see that regular Lasso LARS has a smaller cross-validation error than Lasso LARS using AIC to select α. This is not unexpected as the AIC parameter only looks at the training data. If doing a cross-validation analysis was not feasible, using the AIC modification of Lasso LARS would be recommended. However due to the better results achieved with regular Lasso LARS we choose to not use minimisation of AIC when we fit our magnetic models.

Table 4.2 *Cross-validation of Lasso LARS with AIC minimisation. LS=Least Squares, DDE=Double dipole error*

Non-linear method	Linear regression method	Total fitting error	Total cross-validation error
LS + DDE	Lasso LARS	20.81	87.66
LS	Lasso LARS	20.70	87.93
LS	Lasso LARS AIC	18.99	93.39
LS + DDE	Lasso LARS AIC	19.05	96.75

4.4 Generated models

In Figure 4.4, 4.5, and 4.6 we see a comparison between measurements and the models fit with Lasso LARS for three different vessels from the SIRAMIS data set. We refer to the chosen vessels as *Ship A*, *Ship B*, and *Ship C*. *Ship A* in Figure 4.4 has passed almost directly over the sensor platforms and we consequently have a very good measurement of its signature. The fitted model reproduces this signature well. *Ship B* in Figure 4.5 has passed somewhat to the side of the sensors, and we have measurements of mainly the left side of the ship. We can see that the fitted model extrapolates the field to the right side of the ship in a consistent way. *Ship C* in Figure 4.6 has passed even further to the side of the sensors. With less than half of the magnetic field to work with the fitting routine has recreated a realistic looking signature. In particular we see that the z-component of the field has been recreated with no measurements in the strong part of the field. These three figures showcases the prediction ability of the models fitted with the Lasso LARS algorithm.

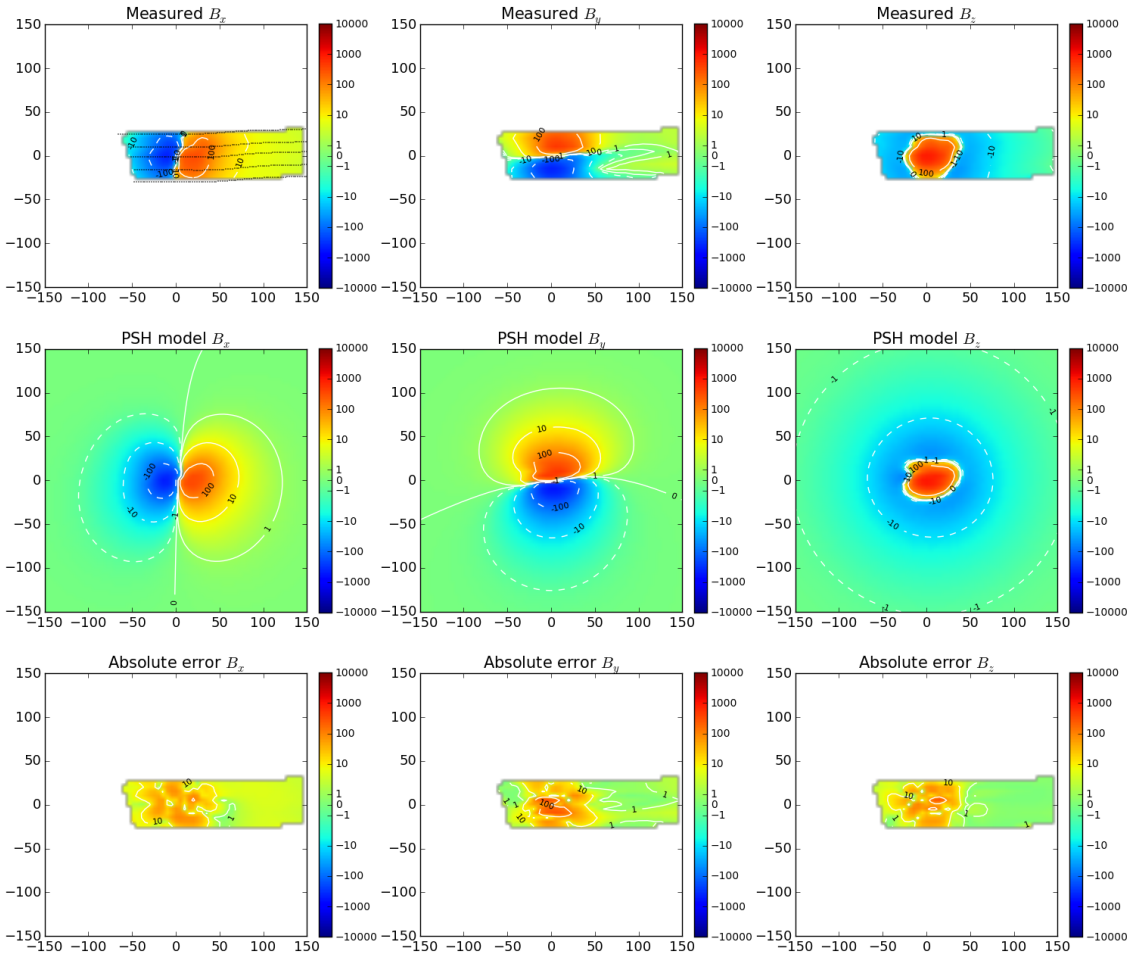


Figure 4.4 *Comparison of measured and modelled x-, y-, and z-components of the magnetic field for Ship A.*

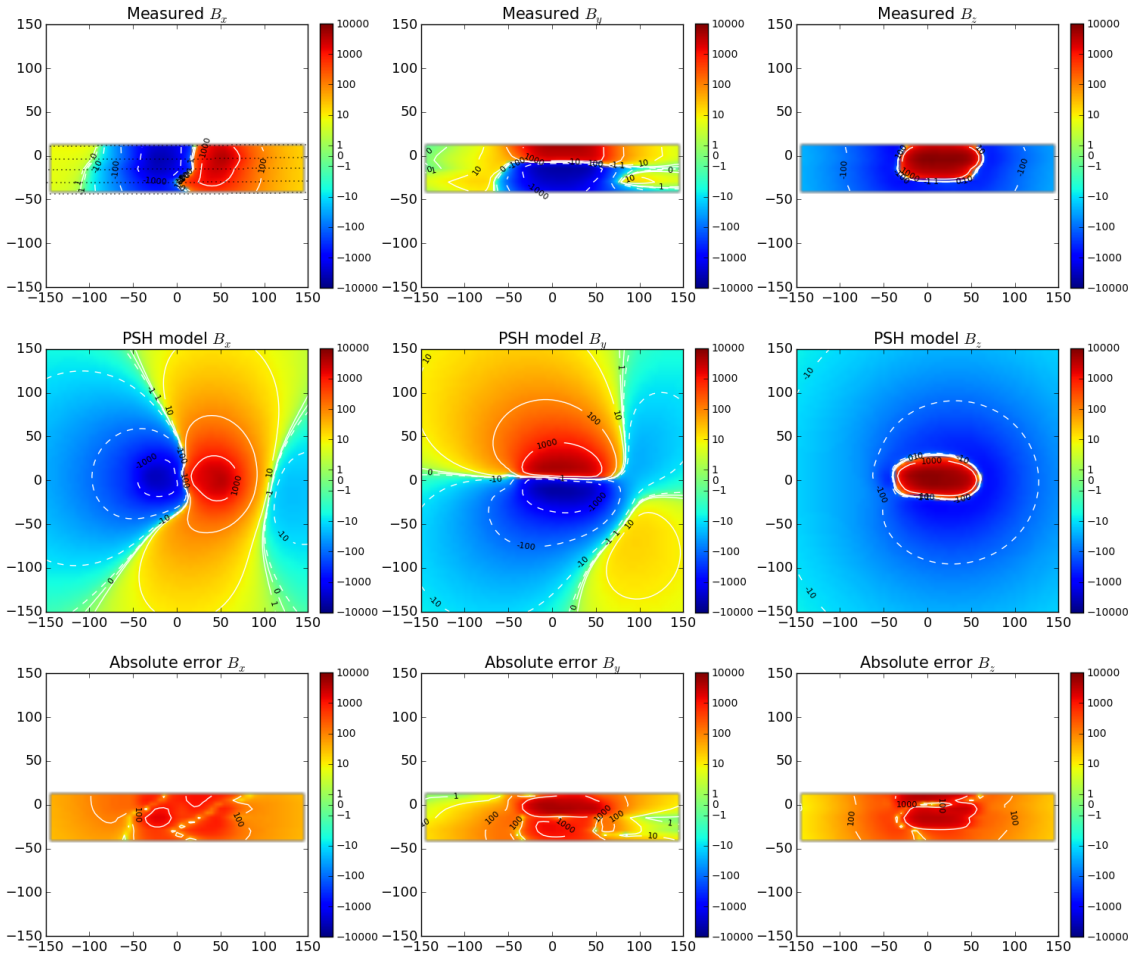


Figure 4.5 *Comparison of measured and modelled x-, y-, and z-components of the magnetic field for Ship B.*

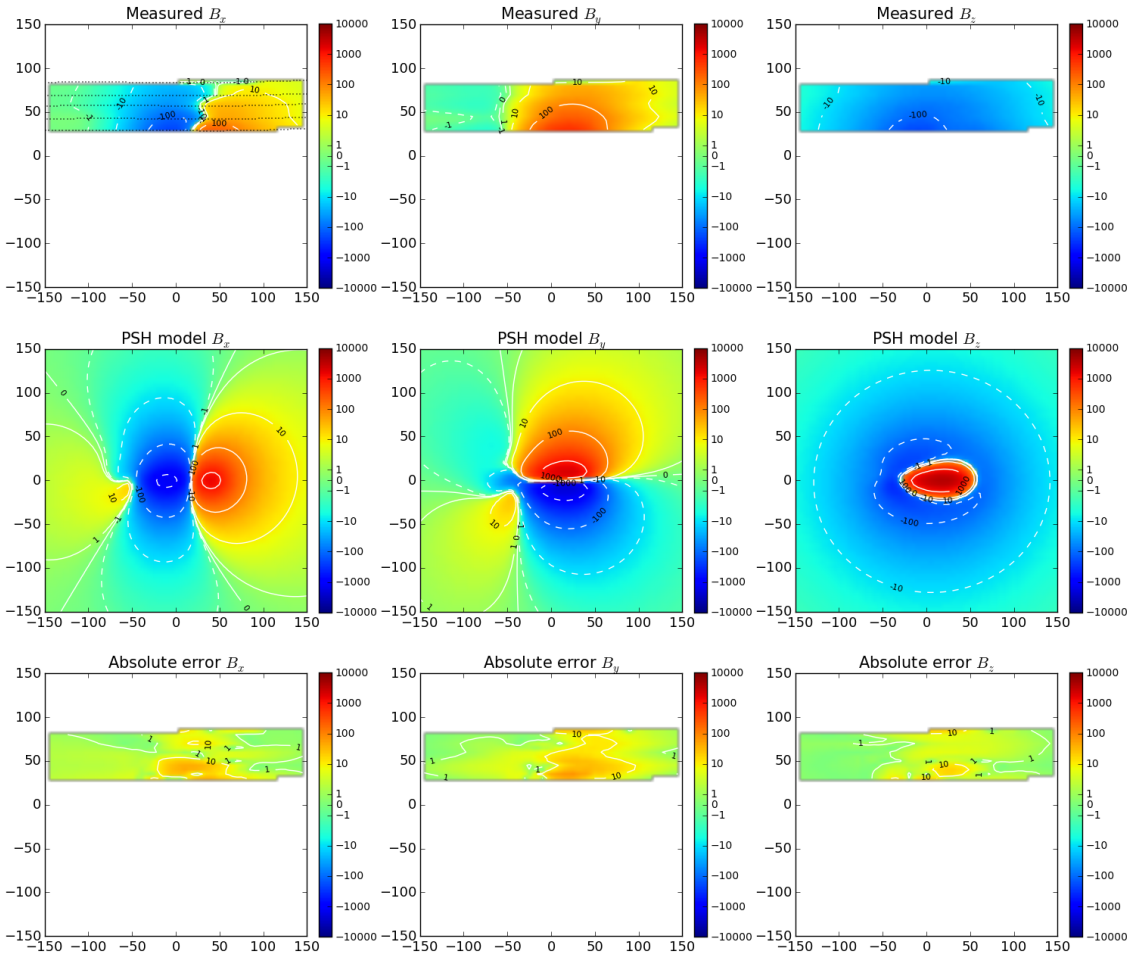


Figure 4.6 *Comparison of measured and modelled x-, y-, and z-components of the magnetic field for Ship C.*

5 Conclusions

A thorough investigation of different methods for fitting the coefficient of the PSH model to the magnetic signatures of civilian merchant ships was performed. The different methods' abilities to produce models with high predictive power were considered and the Lasso LARS method found superior. Other methods, such as the Ridge and ordinary least squares routines overfit the data, resulting in a poor predictive power. Lasso LARS combined model simplicity with a high predictive power, being less sensitive to noise in the training data.

The shape of the coordinate system used for the PSH model was optimised with a non-linear least square method, combined with finding the PSH cofficients with Lasso LARS. A separate regularisation method for the non-linear fit was attempted and found to give better results when combined with linear regression methods that tended to overfit the data, but not with Lasso LARS.

Using Lasso LARS, further studies were performed to tune the method's input parameters to maximise its predictive power and minimise its fitting error. The models generated by our tuned Lasso LARS method were found to reproduce well the measured magnetic signatures. When using measurements which did not capture the entire signature to train the model, it reproduced a realistic looking signature outside the measurements points.

The magnetic ship models generated from the SIRAMIS data set will form the basis for future studies to find a recommendation of performance requirements for new influence sweep systems.

Bibliography

[1] Deliverables from the EDA Project SIRAMIS. P360 Sak 15/02263.

[2] John David Jackson. *Classical Electrodynamics*. Wiley, 3rd edition, 1999.

[3] Alexander V. Kildishev, John A Nyenhuis, and Andrej V. Hetman. Zonal magnetic signatures in spherical and prolate spheroidal analysis. *Marelec*, 1999.

[4] Philip M. Morse and Herman Feshbach. *Methods of Theoretical Physics*. McGraw Hill, 1953.

[5] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2006.

[6] Rungkiet Kamondetdacha, Alexander V. Kildishev, and John A Nyenhuis. Multipole Characterization of a Magnetic Source Using a Truncated SVD. *IEEE Transactions on magnetics*, 40, 2004.

[7] Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12, 1970.

[8] Robert Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, 58, 1996.

[9] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least Angle Regression. *The Annals of Statistics*, 32, 2004.

[10] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on automatic control*, 19, 1974.

[11] Kenneth P. Burnham and David R. Anderson. Understanding AIC and BIC in Model Selection. *SOCIOLOGICAL METHODS & RESEARCH*, 33, 2004.

[12] Stig Asle Synnes, Per Andreas Brodtkorb, and Eugene Lepelaars. Representing the ship magnetic field using prolate spheroidal harmonics - a comparative study of methods. *EMSS*, 2007.

About FFI

The Norwegian Defence Research Establishment (FFI) was founded 11th of April 1946. It is organised as an administrative agency subordinate to the Ministry of Defence.

FFI's MISSION

FFI is the prime institution responsible for defence related research in Norway. Its principal mission is to carry out research and development to meet the requirements of the Armed Forces. FFI has the role of chief adviser to the political and military leadership. In particular, the institute shall focus on aspects of the development in science and technology that can influence our security policy or defence planning.

FFI's VISION

FFI turns knowledge and ideas into an efficient defence.

FFI's CHARACTERISTICS

Creative, daring, broad-minded and responsible.

Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan med særskilte fullmakter underlagt Forsvarsdepartementet.

FFIs FORMÅL

Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

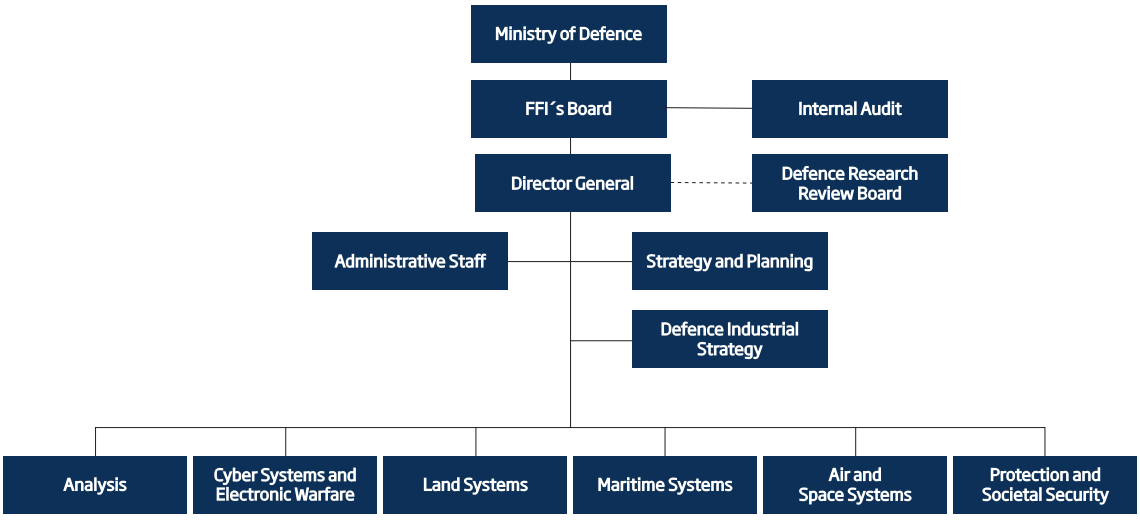
FFIs VISJON

FFI gjør kunnskap og ideer til et effektivt forsvar.

FFIs VERDIER

Skapende, drivende, vidsynt og ansvarlig.

FFI's organisation



Forsvarets forskningsinstitutt
Postboks 25
2027 Kjeller

Besøksadresse:
Instituttveien 20
2007 Kjeller

Telefon: 63 80 70 00
Telefaks: 63 80 71 15
Epost: ffi@ffi.no

Norwegian Defence Research Establishment (FFI)
P.O. Box 25
NO-2027 Kjeller

Office address:
Instituttveien 20
N-2007 Kjeller

Telephone: +47 63 80 70 00
Telefax: +47 63 80 71 15
Email: ffi@ffi.no

FFI Forsvarets forskningsinstitutt