Learning Magnetic Field Distortion Compensation for Robotic Systems

Leif Christensen[1], Mario Krell[2,3] and Frank Kirchner[1,3]

***Abstract*— The work presented in this paper describes the use and evaluation of machine learning techniques like neural networks and support vector regression to learn a model of magnetic field distortions often induced in inertial measurement units using magnetometers by changing currents, postures or configurations of a robotic system. Such a model is needed in order to compensate the local dynamic distortions, especially for complex and confined robotic systems, and to achieve more robust and accurate ambient magnetic field measurements. This is crucial for a wide variety of autonomous navigation purposes from simple heading estimation over standard SLAM approaches to sophisticated magnetic field based localization techniques. The approach was evaluated in a laboratory setup and with a complex robotic system in an outdoor environment.**

I. INTRODUCTION

Current robots are taking a more and more prominent role in the world of today, where they have to cope with increasingly demanding environments. And in correlation with the progress of robotic skills and the advances in the degree of autonomy, expectations are raised and society legitimately demands that robotic systems support humans not only in laboratory environments, but in real world scenarios from everyday situations at home to most challenging and maybe also dangerous tasks.

The area of localization and mapping plays an important role in robotics since decades and has seen huge advances in recent years, nonetheless, there are still open issues especially when leaving laboratory or office-like environments and dealing with long term autonomous robotic operations. Commonly used sensors in such applications often include global navigation satellite systems (GNSS), cameras (stereo, time of flight, monocular, RGB-D, etc.), LIDARs, different kinds of sonars (especially in the underwater domain) as well as a huge variety of other exteroceptive sensors [1, Chapter 6], all of which may fail in certain scenarios depending on the ambient condition of the situation at hand. But one common denominator in nearly all robot navigation applications is the utilization of an inertial measurement unit (IMU), usually consisting of at least accelerometers and gyroscopes to determine orientation. Since gyroscope measurements drift over time, IMUs are often supplemented with a magnetometer to stabilize heading. From seemingly simple absolute heading estimation over standard SLAM approaches to sophisticated magnetic field based localization techniques, these integrated magnetometers play a crucial role by measuring the flux density of the ambient magnetic field.

However, in order to fully exploit the benefits of this convenient combination of an absolute sensor and an almost omnipresent ambient magnetic field, one has to properly deal with the significant dynamic magnetic field distortions caused by ferromagnetic materials or strong electric currents near the magnetometer originating from the robotic system itself. This specifically applies to systems with restricted sensor mounting options far away from distortion sources, for example on very compact robots or autonomous underwater vehicles with pressure housings, but also on complex systems with a lot of moving parts or reconfiguration options like the hybrid legged-wheeled intervention robot *SherpaTT* [2] (Fig. 1) or other forms of complex walking robots.

Fig. 1: The hybrid legged wheeled robot *SherpaTT* with a high amount of degrees of freedom and therefore high amount of magnetic field distortion sources during field trials in the desert of Utah, US

The paper is structured as follows: after this introduction, we discuss magnetic fields and distortions in section II, describe our approach of dynamic distortion model learning in section III and present experimental results in section IV. Finally, we conclude the work presented in section V.

II. MAGNETIC FIELDS AND DISTORTIONS

A. Geomagnetic Field

The most prominent use case for magnetometers (compass applications for heading estimation), depends on measuring the horizontal components of the earth's magnetic field to determine the direction towards the magnetic north pole. Although the pole's location is changing over time and the

[1]German Research Center for Artificial Intelligence, Robotics Innovation Center, Bremen, Germany leif.christensen@dfki.de
[2]University of California Berkeley, International Computer Science Institute, Berkeley, USA
[3]University of Bremen, Robotics Group, Bremen, Germany

Preprint submitted to 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Received February 28, 2017.

magnetic field is significantly locally distorted depending on latitude and longitude of the observer, the declination from magnetic to true north can be computed using analytical models like the *World Magnetic Model* (WMM) [3] or the *International Geomagnetic Reference Field* (IGRF) [4], or numerical models for special situations like field reversals [5]. Apart from this variation coming from the main magnetic field density contributors (the earth core with its geodynamo effect and the earth crust), the earth magnetic field can be considered sufficiently stable in the temporal domain for our purpose. Although there are also variations due to changes in the magnetosphere like SQ-variations or coronary mass ejections, these only have a very small effect, typically less than $250\,\mathrm{nT}$ compared to the strength of the earth magnetic field between $25\,000\,\mathrm{nT}$ to $65\,000\,\mathrm{nT}$, depending on the location [6].

B. System Immanent Distortions

Ferromagnetic materials and strong currents flowing through a wire in the vicinity of the sensor can cause significant distortions of the measured magnetic flux density. These vehicle-immanent distortions are usually classified as hard- and soft-iron effects [7]. Hard-iron effects occur due to the magnetic remanence of nearby material (permanent magnets in motors, magnetized iron or steel) and show a constant offset in every field component. Without distortions, arbitrary magnetic field direction measurements would lie on the surface of a S^2 sphere in $\mathbb{R}^3$. Hard-iron distortions however lead to a shift of the center of the sphere from the origin (one-cycle error, see Fig. 2) and can be modeled as a 3-component bias vector $b_{hi} = \begin{pmatrix} x_{hi} & y_{hi} & z_{hi} \end{pmatrix}^T$.

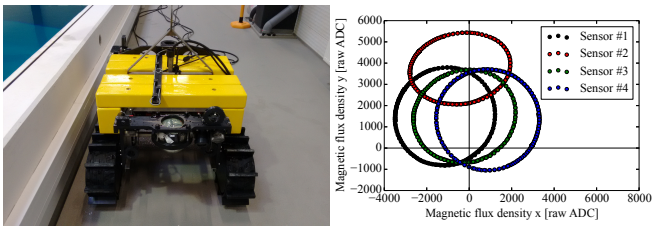

Fig. 2: Hard-iron distortion and misalignment of multiple magnetometers on robotic crawler *Wally* projected onto the xy-plane leading to off-center effects, depending on their mounting position

Strong currents flowing through wires near the magnetometers also lead to hard-iron effects, but are usually non-static, while soft-iron effects distort the magnetic field by providing a path of lower impedance while an external field is applied to the ferromagnetic compound. This induces magnetism depending on the orientation of the material with respect to the applied field (two-cycle error). As such, soft-iron effects lead to deforming the sphere to a 3D ellipsoid, but retaining the origin. The classic approach to deal with these distortions is to trying to mount the magnetometer as far away from dynamic distortions sources as possible to minimize dynamic distortions and for the static part compute calibrated parameters $\hat{x}, \hat{y}, \hat{z}$ from raw sensor readings x, y, z by finding the values for the following equation in an a priori one-shot calibration procedure:

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} = M_{align} \cdot \begin{pmatrix} sc_x & 0 & 0 \\ 0 & sc_y & 0 \\ 0 & 0 & sc_z \end{pmatrix} \cdot M_{si} \cdot \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} - b_{hi} \right)$$

with the misalignment matrix M_{align}, a diagonal scale/conversion matrix, soft iron distortion matrix M_{si} and hard-iron offset vector b_{hi}.

Depending on the severity of the system-induced and dynamically changing field distortions in the vicinity of the sensor, a priori calibration techniques can correct the measurements only to a certain point and may fail completely on systems with moving ferromagnetic parts.

III. DYNAMIC DISTORTION MODEL LEARNING

System-immanent dynamic disturbances are a strong contributor to distortions of the otherwise evenly distributed ambient magnetic field, especially in complex and confined robotic systems, as discussed in section II-B. To be able to use the ambient magnetic field for orientation or localization purposes in the first place, the dynamic distortions have to be compensated for.

The rationale behind the chosen approach is that in most recent robotic systems there is a huge amount of proprioceptive sensor data available at runtime that can help to deduct the magnetic field distortions emanating from the system. For example, we often have means to measure the actual currents flowing through wires or torques applied to the motors. Apart from full reflex-driven robots, we will most of the time have quite accurate data on the relative position of extremities and appendages of the robotic system (e.g. in legged robots) as well as current state information (e.g. attached payloads or robot configuration) in reconfigurable robots. For systems with restricted access to the proprioceptive sensor data (often the case with commercial systems), an approach for dynamic distortion filtering using a distributed magnetometer array is described in [8].

While it is possible to facilitate some simplifications and model certain distortions as bar magnets, the sheer amount of contributing and intertwined magnetic field distortion sources in the systems in consideration almost always renders the formulation of an analytical solution impossible.

The approach in this paper is therefore to learn a function f of the resulting and superimposed distortions at the point of the magnetometer sensor from the proprioceptive sensor data of a robotic system (Fig. 3). Since we are not classifying but try to establish a function and the target to be learned is in our case a 3D offset vector, the problem falls in the class of multi-target function regression. To evaluate our approach with different regression techniques and meta-parameter optimization, we used the *scikit-learn* framework [9] in combination with the robotic framework *Rock* [10].

A. Support Vector Regressor

The basic concept of support vector regression (SVR) is to learn a linear function such that the given training data lies

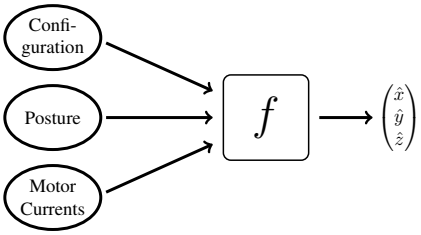

Fig. 3: Multi target function regression approach from proprioceptive robot data

within an epsilon-tube around this function, i.e., distances between the learned function and the given values in the training data are less than epsilon [11]. Larger errors are linearly penalized. Furthermore, a regularization term is preferring smooth functions with small weights. The weighting between error (loss) and regularization (small weights) is done with a regularization constant C. For modeling complex functions, the kernel trick is applied [12], [13]. Using the very common radial basis function kernel ($e^{-\gamma||x_i - x_j||^2}$, RBF), an additional hyperparameter γ has to be chosen.

With the SVR approach, we have to train one support vector machine for each of the three dimensions of the target offset vector, which somewhat neglects the fact, that these 3 components are inherently coupled, because they describe a magnetic flux density vector, incorporating field orientation and strength.

B. Multi Layer Perceptron Regressor

A neural network in the form of a multi layer perceptron regressor (MLP) can innately represent coupling between components and directly be trained for multi target regression in contrast to a SVR. In our case, the input layer represents the different commands and sensor inputs that can influence the magnetic field measurement and the output is a representation of the different components of the magnetic field like directions and strength. The core components of an MLP are perceptrons that linearly weight the different inputs and apply a gating/activation function afterwards [14]. Each layer of an MLP consists of several perceptrons that are not connected to each other but to all perceptrons in the preceding and the follow-up layer. For learning the weights of the single perceptrons, numerous optimization strategies can be used that are often able to handle huge amounts of data [15].

IV. Experiments

In order to evaluate the overall compensation performance of our approach, e.g. how well the learned model can keep the directional component of the magnetometer measurements stable in the presence of local dynamic magnetic field distortions, we conducted experiments with an artificial distortion turntable setup as well as with the hybrid wheeled-legged robot *SherpaTT*. The tests were performed in a very noise free environment in the Mars-like desert of Utah, US, during an extended field trials period with a team of heterogeneous robotic systems [2], [16].

To compare the compensation performance, we are using two different probabilistic distributions. This is because the 3 components measured by the magnetometer are not isolated, but represent the strength and direction of the local magnetic flux density field. For example sometimes even strong distortions result in a strong change of the magnetic field strength, whereas the directional component is rather steady and vice versa. We model the strength component as a Gaussian distribution of the $L2$-norm with mean μ_{st} and standard deviation σ.

To model the three-dimensional direction component, we use the *von Mises Fisher* (vMF) distribution [17], originating from directional statistics, analog to circular wrap-around distributions in the one-dimensional case. The vMF distribution is defined on the S^{p-1}-dimensional sphere in $\mathbb{R}^p$. The probability density function of a vMF distribution on S^2 is given by

$$p(\mathbf{x_i}|\mu_{dir}, \kappa) = \frac{\kappa}{4\pi \sinh \kappa}$$

with mean direction μ_{dir} and concentration parameter κ for a unit direction vector $\mathbf{x}$. $\kappa = 0$ means uniform distribution, while it is more concentrated with higher κ (see Fig. 4, in our application, higher kappa means better directional compensation).

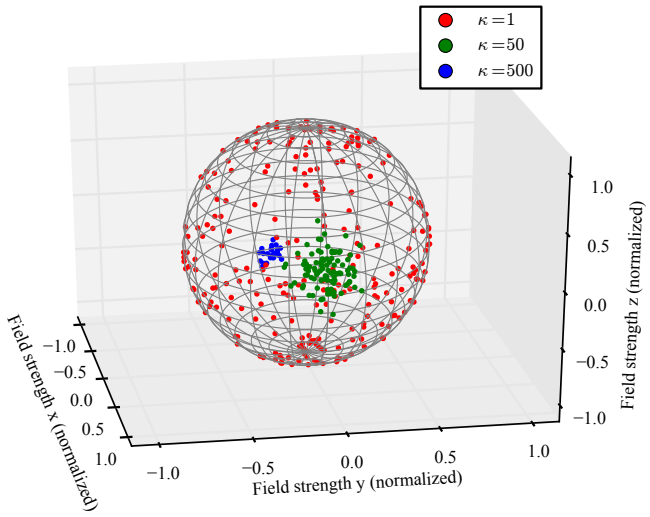

Fig. 4: Samples from three different vMF-distributions on S^2 with different mean and concentration parameter κ

We approximate μ_{dir} as

$$\hat{\mu}_{dir} = \frac{\mathbf{r}}{\|\mathbf{r}\|} = \frac{\sum_{i=1}^{n} \mathbf{x_i}}{\|\sum_{i=1}^{n} \mathbf{x_i}\|}$$

and κ according to [18] and as proposed for small dimensions by [19] as

$$\hat{\kappa} = \frac{\bar{r}d - \bar{r}^3}{1 - \bar{r}^2} \text{ with } \bar{r} = \frac{\|\mathbf{r}\|}{n}$$

A. Magnetic Field Distortion Turntable

In order to test the approach with very defined and separable distortion sources, we created an artificial turntable

Preprint submitted to 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Received February 28, 2017.

setup as shown in Fig. 5. To resemble hard-iron distortions a neodym magnet was mounted to a lever-arm (1) moving $90°$ arc-wise towards the magnetometer mounted in the center of the turntable. To emulate soft-iron distortions, a $1\,\mathrm{kg}$ $99.9\,\%$ pure iron block (2) was moved $6\,\mathrm{cm}$ linearly towards and away from the sensor. And finally, to simulate electromagnetic distortions from motor supply currents, a $6\,\mathrm{mm}^2$ wire (3) was fixated close to the sensor (4). The material of the setup was chosen in order to minimize magnetic distortion sources different from those named above. Also every piece of equipment brought to the experimental site was validated to not interfere with the experimental setup.

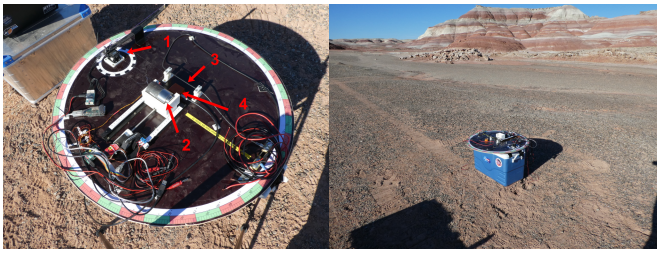

Fig. 5: Magnetic distortion turntable experiment in the desert of Utah, US

During data recording, every distortion source activity was activated individually and then simultaneously. In each trial, the hard iron source lever was moved $90°$ twice in an arc towards the sensor. The soft iron source was moved once towards the sensor and then back, while the current was raised from $0\,\mathrm{A}$ to $20\,\mathrm{A}$ and back to $0\,\mathrm{A}$ to simulate electromagnetic disturbances from supply lines. After each trial, the turntable was rotated by $20°$ to eventually achieve a full circle. Whereas the hard iron source, as expected, had the strongest impact on the magnetic field when coming near to the sensor, all other distortion sources were clearly also superimposing the ambient magnetic field (see Fig. 6).

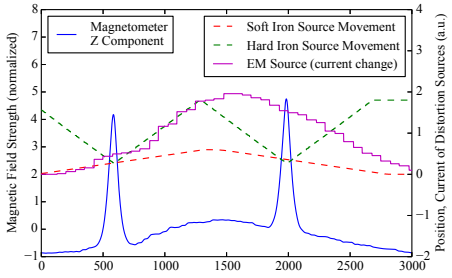

Fig. 6: Magnetometer z component plotted against the superimposing distortion source activities (turntable trials, heading $80°$)

Before applying the different machine learning (ML) techniques, we applied a second order Butterworth filter to filter out high frequency noise. For evaluation, the turntable dataset was randomly split into a training set ($60\,\%$) and a test set ($40\,\%$). We then applied a k-fold cross-validation grid search with 5 splits on the training set to prevent leakage of knowledge about the test set into the model during hyperparameter tuning. We used a grid over the ranges α: 1×10^{-3} to 1×10^{-7}, number of hidden layers between 1 to 3 with 5 to 100 perceptrons per layer, solvers *Adam* [20] and *LBFGS* [21], and activation functions *logistic*, *relu*, and *tanh*. Training our MLP regressor with 4 inputs (3 distortion sources plus heading) and 3 outputs, we achieved best performance using 2 hidden layers of size $[10, 20]$, an α value of 1×10^{-4}, *relu* as activation function and *LBFGS* as a solver. The prediction and compensation quality on the turntable data set is shown in Fig. 7.

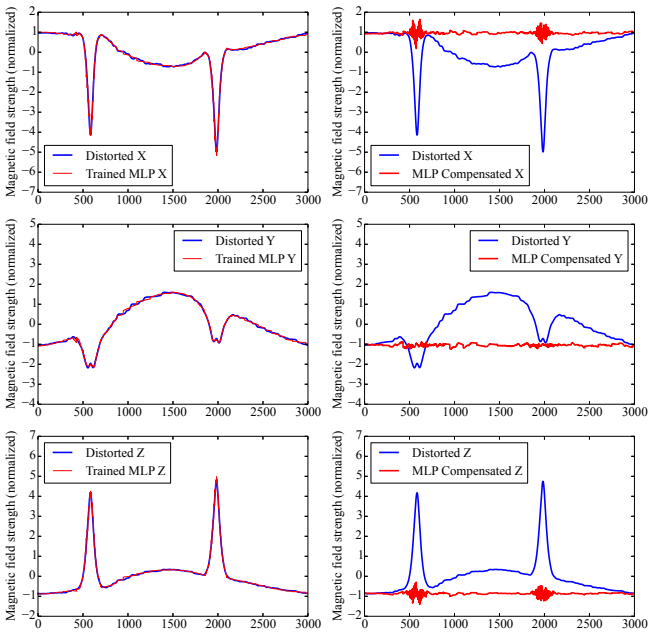

Fig. 7: Component-wise MLP prediction (left) and compensation (right) of magnetic field distortions (turntable trials, heading $80°$)

The MLP Regressor with the parameters presented above achieved an R^2-score [22] of 0.986. Finally, using the MLP Regressor to predict the magnetic field distortions at the point of the magnetometer sensor given only the values of the activity for the various distortion sources, we achieved a significant reduction of the deviation due to dynamic distortions in the direction component, reflected in an increase in the κ concentration parameter from 0.86 to 618.2. See Fig. 7 for a component wise comparison and Fig. 8 for a 3D directional scatter plot.

B. SherpaTT dataset and evaluation

Apart from the turntable experiment, we also evaluated our approach by a series of experiments with the complex hybrid wheeled legged robot *SherpaTT*, to analyze to what extend it is transferable to real robotic systems. The idea, as in the turntable setup, was to try to keep the orientation of the magnetometer stable in the ambient magnetic field and then activate as many measurable distortion sources as possible, both solitary and in combination, and record the induced vector field deviation from the sensor baseline. We did this by repeating a sequence of leg movements of the robot, first trying to cover most of the robot's DOF workspace, and second varying the single joint ranges while maintaining the

Preprint submitted to 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Received February 28, 2017.

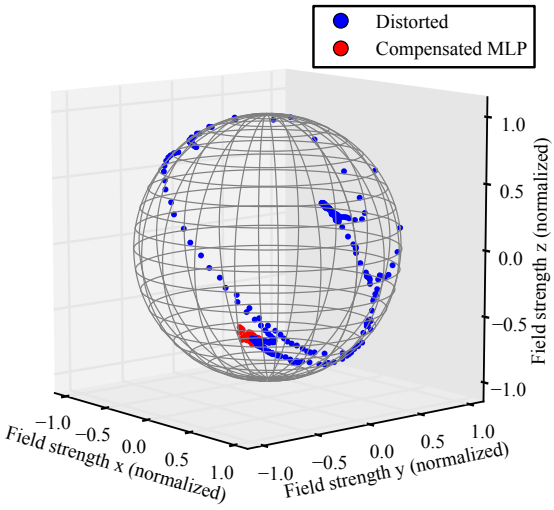

Fig. 8: 3D scatter plot of dynamically distorted vs. MLP compensated directions (turntable trials, heading 80°). Every dot represents the direction of a magnetic field direction measurement. Undistorted measurements would stay on the same spot on the sphere's surface.

central body pose. Furthermore, we generated strong changes in the supply current. Since soft-iron types of distortion depend on the orientation in the ambient magnetic field, we recorded the data set in 45° steps, covering a full 360° spot turn circle. Extra care was taken to prevent external disturbances during the data gathering.

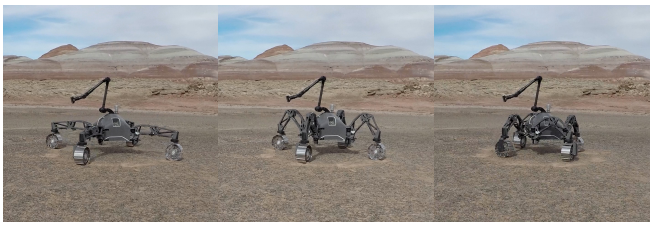

Fig. 9: *SherpaTT* during magnetic field distortion data set gathering

Whereas we had few distinct and strong distortion sources in the turntable experiment (IV-A), in the experimental setup using the robot *SherpaTT* we had a multitude of permanent magnets moving around in each actuated robotic leg joint and the manipulator as well as multiple power supply lines in varying distance from the sensor for the four legs and the robotic arm. In general, we observed deviations of the magnetic field measurements orders of magnitude smaller than the deviations that occurred during the turntable experiment. This was expected, since the magnetometer was positioned further away from possible distortion sources inside the robot's main body housing than in the turntable experimental setup, where we intentionally moved or placed the distortion sources verly close to the sensor.

This time, we also trained a SVR with two different kernels (linear and RBF) for comparison with the MLP regressor. The MLP had an input layer of size 25 (5 joint positions and one supply line per leg plus heading) and again 3 outputs representing the magnetic field offset vector.

To obtain the best parameters for the SVR and MLP, we again applied a grid search over the the same ranges as in the turntable experiment for MLP and from 0.1 to 25 for the SVR C parameter. For the SVR γ parameter, we used the *auto* functionality of `sklearn.svm.SVR`. As with the turntable data set, the *SherpaTT* deviation dataset was randomly split into a training set (60 %) and a test set (40 %). We again applied a k-fold with 5 splits on the training set for hyperparameter tuning. We achieved best performance using 3 hidden layers of size $[100, 50, 25]$, an α value of 1×10^{-4}, *tanh* as activation function and again *LBFGS* as a solver (which is common for small training datasets) for MLP and $C = 1.5$ for SVR with RBF kernel. The resulting predictions on the test data are shown in Fig. 10 for SVR and in Fig. 11 for MLP.

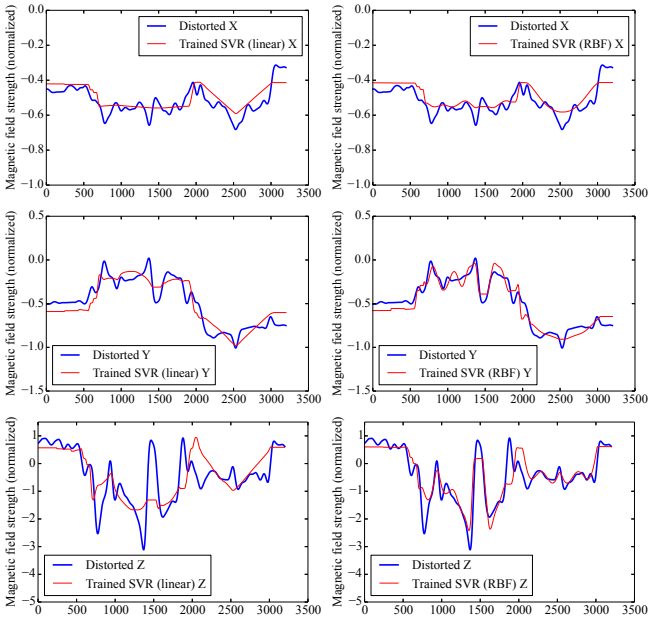

Fig. 10: Component-wise SVR prediction of magnetic field distortions with linear (left) and RBF (right) kernel (*SherpaTT* trials, heading 180°)

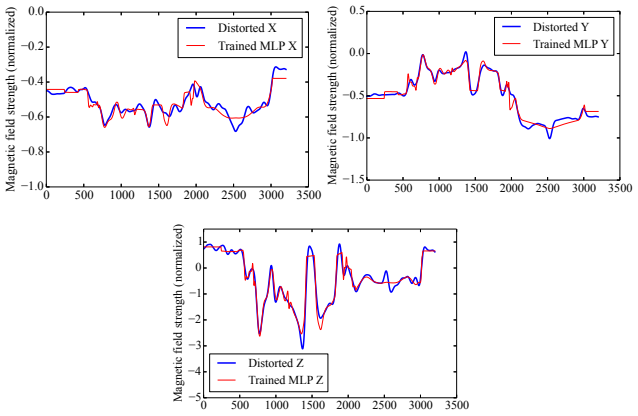

Fig. 11: Component-wise MLP prediction of magnetic field distortions (*SherpaTT* trials, heading 180°)

Preprint submitted to 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Received February 28, 2017.

The multi target MLP model with its optimized meta parameters was able to achieve a much better fit on the test data set with R^2 training scores of 0.96 compared to the single component SVR models with 0.60, 0.84, 0.46 (X, Y, Z linear kernel) and 0.63, 0.91, 0.80 (X, Y, Z RBF kernel). This is also reflected in the compensation quality of the directional component of the magnetic field: the SVR based compensation was not able to stabilize the direction to the same extent as the MLP based compensation, with the respective compensation parameters $\kappa = 8.21$ for SVR with linear kernel, $\kappa = 15.91$ for SVR with RBF kernel and $\kappa = 114.50$ for MLP compared to the distorted directions with concentration $\kappa = 3.64$ (Fig. 12). This indicates that for modeling the magnetic field, simple interpolation approaches like SVR with linear or RBF kernels are insufficient and more complex functions like those represented by a 3-layer MLP are required.

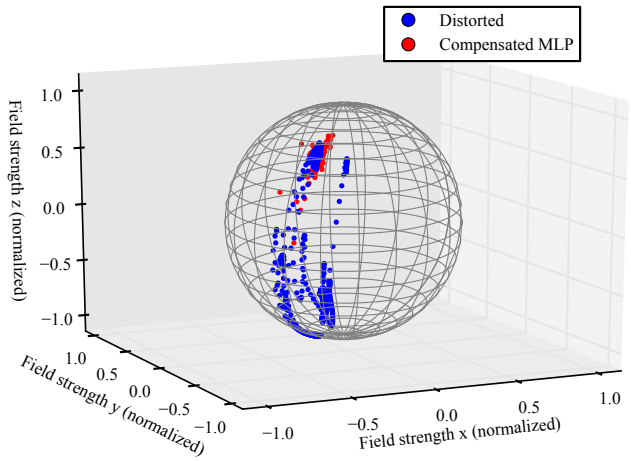

Fig. 12: 3D scatter plot of dynamically distorted and MLP compensated directions (*SherpaTT* trials, heading $180°$)

V. CONCLUSIONS

In this paper, we presented and discussed a ML approach to model magnetic field distortions arising from dynamic changes in complex robotic systems. We found that, compared to SVR, MLP regressors with LBFGS solvers are especially capable of predicting the magnetometer deviations and we presented promising results of direction compensation based on such an approach not only in a laboratory setup, but also on a complex real robot. While the MLP fits the data quite well, in future work we would like to investigate more the robustness and transferability of our approach in combination with sparse training input data. It would also be interesting to derive models that can provide more insight about the distortion contribution of the different input sources.

We think having a ML-based tool to establish magnetic field distortion models as presented in this work may become relevant in a lot of upcoming real world scenarios in robotics, since it widens the usability of magnetometers as one of the core sensors in many navigation applications.

ACKNOWLEDGMENT

The authors would like to thank all colleagues at the DFKI Robotics Innovation Center for their support and feedback, and especially Florian Cordes, Roland Sonsalla, Thomas Roehr, Tobias Stark and Steffen Planthaber of the Field Trials Utah team. The research leading to this result has received funding from the European Union FP7 program (grant No. 312762) and the German Federal Ministry of Economics and Technology (BMWi), projects TransTerrA and FT Utah (grant No. 50RA1301, 50RA1621).

REFERENCES

[1] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.
[2] F. Cordes and A. Babu, "SherpaTT: A versatile hybrid wheeled-leg rover," in *Proceedings of the 13th International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS 2016)*, 2016.
[3] Chulliat, A., S. Macmillan, P. Alken, C. Beggan, M. Nair, B. Hamilton, A. Woods, V. Ridley, S. Maus, and A. Thomson, "The US/UK world magnetic model for 2015-2020," NOAA National Geophysical Data Center, Boulder, CO, Tech. Rep., 2015.
[4] C. C. Finlay et al., "International geomagnetic reference field: the eleventh generation," *Geophysical Journal International*, vol. 183, pp. 1216–1230, 2010.
[5] G. A. Glatzmaier and P. H. Roberts, "A three-dimensional self-consistent computer simulation of a geomagnetic field reversal," *Nature*, vol. 377, no. 6546, pp. 203–209, 1995.
[6] M. McElhinny and P. L. McFadden, *The magnetic field of the earth: paleomagnetism, the core, and the deep mantle*. Academic Press, 1998, vol. 63.
[7] M. Caruso, "Applications of magnetic sensors for low cost compass systems," in *Position Location and Navigation Symposium, IEEE 2000*, 2000, pp. 177–184.
[8] L. Christensen, C. Gaudig, and F. Kirchner, "Distortion-robust distributed magnetometer for underwater pose estimation in confined UUVs," in *Proceedings of MTS IEEE OCEANS '15*. Washington DC, USA: IEEE, 2015, pp. 1–8.
[9] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[10] "ROCK, the Robot Construction Kit," http://www.rock-robotics.org.
[11] V. Vapnik, "The nature of statistical learning theory," *Springer*, 2000.
[12] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.
[13] C.-C. Chang and C.-J. Lin, "LIBSVM," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.
[14] G. E. Hinton, "Connectionist learning procedures," *Artificial Intelligence*, vol. 40, no. 1-3, pp. 185–234, 1989.
[15] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Aistats*, vol. 9, 2010, pp. 249–256.
[16] R. U. Sonsalla, J. B. Akpo, and F. Kirchner, "Coyote III: Development of a modular and highly mobile micro rover," in *Proceedings of the 13th Symposium on Advanced Space Technologies in Robotics and Automation (ASTRA-2015)*, Noordwijk, The Netherlands, 2015.
[17] R. A. Fisher, "Dispersion on a sphere," in *Proc. Roy. Soc. London*, ser. A, vol. 217, 1953, pp. 295–305.
[18] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra, "Clustering on the unit hypersphere using von Mises-Fisher distributions," in *Journal of Machine Learning Research*, 2005, pp. 1345–1382.
[19] S. Sra, "A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of $I_s(x)$," *Computational Statistics*, vol. 27, no. 1, pp. 177–190, 2012.
[20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014. [Online]. Available: http://arxiv.org/abs/1412.6980
[21] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical programming*, vol. 45, no. 1, pp. 503–528, 1989.
[22] R. Anderson-Sprecher, "Model comparisons and R^2," *The American Statistician*, vol. 48, no. 2, pp. 113–117, 1994.