Fast recovery of magnetic vacuum configuration of WEGA stellarator with error field effects

A.Sengupta, J.Geiger[1], M.Otte, A.Werner

Max-Planck-Institut für Plasmaphysik, Euratom Association, Greifswald, Germany

With a background of having obtained positive results with Function Parametrization (FP) applied to stellarator configurations, the technique was used once again for recovering the vacuum magnetic field configurations of the WEGA stellarator including the main symmetry-breaking magnetic islands. A classical stellarator of type $l = 2$, WEGA has an inherent $n = 1$ (leading order) field perturbation responsible for these islands. The perturbation is assumed to be generated by a misalignment between the centres of the toroidal and helical field generating coil systems. These $n = 1$-periodic WEGA configurations, displaying no stellarator symmetry, were numerically generated around the experimental boundaries and analysed with FP. For the first time FP models with 4th order polynomials and non-linear regressions with rational functions were needed to parametrize the physical state of the configurations. Modelling of the widths of the magnetic islands was challenging, however. The FP functions are in the process of being implemented to run with the WEGA control system.

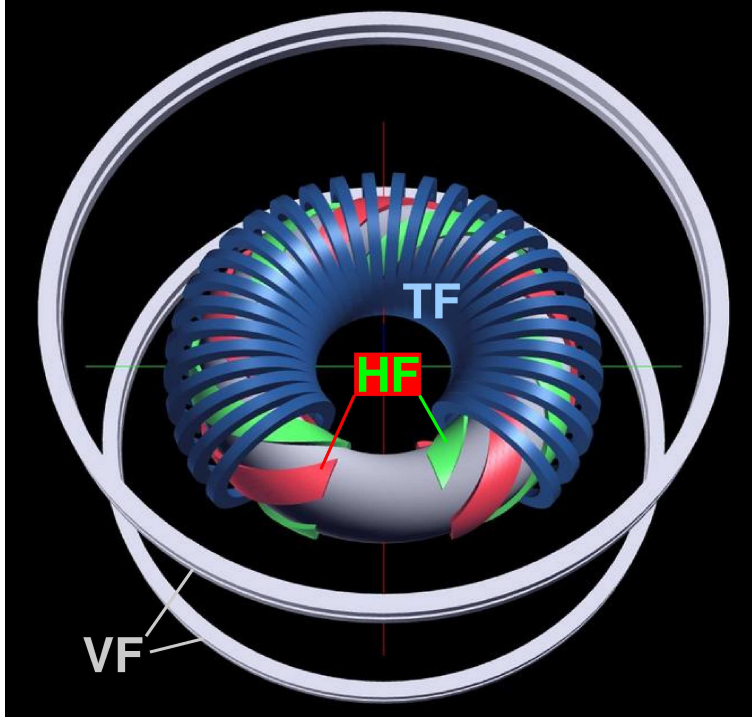[1]Corresponding author. e-mail: joachim.geiger@ipp.mpg.de

Figure 1: The WEGA stellarator showing the external coils. TF: toroidal field coils. HF: helical field coils. VF: vertical field coils.

1. INTRODUCTION

Fast recovery of magnetic configurations is a crucial issue for all fusion devices. Being inherently steady state devices stellarators need rapid and reliable techniques, which would monitor the evolution of different physics parameters in real time, to be implemented with them.

The WEGA stellarator [1] (Fig. 1) is in operation since July 2001 at the Greifswald branch of Max-Planck-Institut für Plasmaphysik in Germany. A classical stellarator intended primarily for educational purpose, it has a major radius $R = 72$ cm, a plasma radius $a_{eff} \leq 11.5$ cm and a planar and circular magnetic axis. 40 toroidal field (TF) coils and 2 pairs of helical field (HF) coils generate an $l = 2$ type stellarator configuration with a five-fold toroidal periodicity and stellarator symmetry. The HF coils produce the rotational transform ι required for plasma confinement with an usually-operated range between 0.1 and 1.0, limited by the plasma size which becomes too small to be of use for large values of iota (typically for axis values of $\iota_{ax} > 0.8$). Two pairs of vertical field (VF) coils of the Helmholtz type, one pair above and the other below the torus, provide control over the magnetic axis position. The shapes of the flux surfaces, in their toroidal variation, are rotating ellipses which are characteristics of $l = 2$ helical windings. The plasma (resonant) start-up is by means of a 2.45 GHz Electron Cyclotron Resonance Heating (ECRH) system at a magnetic field strength of ≤ 87.5 mT [2]. With a recently implemented 28 GHz ECRH the machine is operated at $\sim$0.5 T. For the high-field operation, technical constraints such as the cooling system for the HF coils and the pulse length of the discharges further limit the operational ι_{ax}-range to 0.2 - 0.5. The plasma boundary is defined either by the ECRH antennae or the inner wall of the vacuum vessel acting as material limiters.

However, real configurations do not enjoy the periodicity and the stellarator symmetry of the ideal

coil system as described above, but suffer from a breakdown of both these features by error fields. Error fields are perturbations on the symmetric and periodic magnetic configuration and arise from, e.g., inaccuracies in manufacturing and assembling of the external coils, deformations in the coils during operation, etc. In case of WEGA, error fields are due to an assumed horizontal misalignment in the mm-range between the TF and the HF coil systems [1]. The periodicity-breaking effect was observed in flux surface measurements from the non-natural islands at low-order values of the rotational transform, e.g. the 1/4 island instead of the natural 5/20 islands. The extent of the misalignment was inferred from a comparison of experimentally measured flux surfaces and rotational transform profiles with numerically calculated ones. A 4 mm misalignment reproduces the observed islands and their sizes in the best way. As a result of the break down of the stellarator symmetry, the toroidal periodicity has a dominant $n = 1$ mode. In flux surface measurements islands belonging to higher modes ($n \geq 2$) are observed, but with smaller island widths due to the higher poloidal mode numbers needed to be resonant with the corresponding rational iota-value. Nevertheless, this gives rise to unique features for the entire torus.

Encouraging results were obtained with Function Parametrization (FP) for W7-X parameter recovery where cubic polynomials were used to model vacuum configurations [3] in terms of external coil current ratios (CCR), and the finite-beta case [4] in terms of CCR, plasma pressure and toroidal plasma currents. Since error field effects were always neglected, it was now decided to test the applicability of the method on a smaller device whose magnetic configurations included error field effects. Because WEGA is a stellarator in operation, the main motivation for this study was that the analysed real magnetic configurations can be implemented as FP-functions into the control system of WEGA to work in real time in same way as planned for W7-X.

In this paper an analysis of the vacuum configurations of WEGA is presented, including the non-periodic, symmetry-breaking magnetic islands having their toroidal periodicity equal to the perturbed value of $n = 1$. Recovering the important magnetic islands in the configuration formed a crucial part of this work. The following section briefly discusses the basic principles of FP, while Section 3 describes the nature and the size of the database that was analysed to set up the FP model. The detailed results of the statistical analysis are in Section 4, first for the scalar, or locally determined, parameters, and then for the profile parameters, meaning those varying with the flux surfaces.

2. Principles of FP

The basic principle of FP [3 - 7] consists in getting a simple representation of a certain dependent (or response) variable Y, as a function of a set of independent (or predictor, regressor) variables, resulting from a statistical analysis on a large dataset that contains all these variables. For convenience, let the set of predictor variables define a vector $\vec{x}$. The ranges of variables over what is also called the training dataset should encompass those expected in the experiment as these statistical models learn the trends of variation within the data and are usually poor extrapolators if trend changes are present outside of the training dataset, but also especially if non-linear models are used.

A typical FP model has the form

$$y = F(\vec{x}) + \epsilon \tag{1}$$

where the function F is an estimate of the dependency of the true value y of the response variable Y on $\vec{x}$, and ϵ is a random error term which shows that this representation is only an approximate one. Therefore one may write

$$\tilde{y} = F(\vec{x}) \tag{2}$$

where $\tilde{y}$ is the FP-estimated value of Y.

The components of $\vec{x}$ should be statistically uncorrelated among themselves and should not have widely varying standard deviations, to make the problem well-conditioned. If they are correlated, the raw variables are subjected to a coordinate transformation (e.g., Principal Component Transformation PCT [8]) such that the transformed variables, which will replace the raw variables as the components of $\vec{x}$ in equation (1), are mutually orthogonal in the new space. If their standard deviations differ by orders of magnitude, the raw variables need to be standardized (i.e., normalized to unit standard deviation) before being analyzed.

Also known as "regression" in statistics, the above function is set up using the principles of least squares whereby the coefficients of F are estimated from the minimisation of a mean squared error (equation (3) below) when $\tilde{y}$ is compared with y:

$$\langle \epsilon^2 \rangle = \frac{1}{N} \sum_{\alpha=1}^{N} [y^{(\alpha)} - \tilde{y}^{(\alpha)}]^2 \tag{3}$$

Here α runs over the N observations in the database. Different forms of the function F result in different estimates of y. An error analysis shows the kind of function that best fits y to the data.

In order that the FP model is robust enough to be used later with data not contained in the training set, two important conditions must be satisfied. First, the new data must come from within the same configuration space, and be generated using the same criteria, as the training data. This is important to ensure that the models are not forced to extrapolate beyond the boundaries of the training data. Secondly, there should be a sufficient difference between the numbers of training data points and the model parameters being estimated. This difference is called the residual degrees of freedom after the model is set up, and determines the generalization capability of statistical models. Therefore, the quality of the fit of equation (1) is tested by another error analysis on an independent but known subset of the training dataset, called the test dataset.

These time-consuming offline steps precede the ultimate application of equation (1), when new data for $\vec{x}$ are fed in to calculate unknown y_i. This process is very fast as it simply involves evaluating F.

Finally, a regression is termed as linear (non-linear) depending on whether the coefficients of F are linear (non-linear) in the equation for y_i.

3. THE DATABASE OF WEGA CONFIGURATIONS

As described in the preceding section, setting up of an FP model relies on analysing a (training) dataset showing the trends of variation of all relevant variables that the model is to approximate. This dataset, generated by conventional computer codes, needs to be produced carefully in order not to introduce numerical errors. However, for some variables their numerical determination is connected with an uncertainty which cannot be specified in any case, e.g. for variables connected to the configuration

boundary and to the island parameters. Another such parameter is the iota-value, because it is calculated by tracing a finite length of a field line and also due to the algorithm used. Nevertheless, as long as the errors are somewhat random (comparable to noise) or low enough for the required accuracy (in case of iota) and do not destroy the underlying trend we can still regress the trend if we have enough training data. The underlying uncertainty will, however, show up in a lower limit of the FP-recovery error, below which the error cannot be reduced even if the models are refined. Therefore, in the error analyses to determine the quality of the FP model we are looking for the saturation effect that tells us whether the numerical error is negligible with respect to the resulting reconstruction error. In the next section this will be examined further.

Table 1: Database summary statistics of WEGA parameters.				
Parameter	mean	spread σ	minimum	maximum
ι_{ax}	0.31	0.215	0.017	0.837
ι_b	0.33	0.218	0.013	0.867
a_{eff} (cm)	8.91	4.40	1.60	18.95
R_{ax} (cm)	70.58	3.95	60.03	83.50
B_{ax} (T)	0.30	0.018	0.25	0.35
$r_{is}^{(1/3)}$ (cm)	8.09	2.44	1.73	11.91
$w_{is}^{(1/3)}$ (cm)	2.05	1.15	0.74	5.78
$r_{is}^{(1/4)}$ (cm)	8.58	2.06	2.73	12.26
$w_{is}^{(1/4)}$ (cm)	1.08	0.45	0.31	3.94
$r_{is}^{(1/5)}$ (cm)	7.64	2.57	1.71	12.65
$w_{is}^{(1/5)}$ (cm)	0.59	0.41	0.20	2.25

The physical parameters describing the WEGA configurations, which are listed as response variables for the regression in table 1, contain: the rotational transform on the magnetic axis (ι_{ax}) and at the boundary (ι_b), the axis position R_{ax}, the on-axis magnetic field strength B_{ax}, the position $r_{is}^{(n/m)}$ and the width $w_{is}^{(n/m)}$ of the important non-periodic, symmetry-breaking magnetic islands in the configuration (in particular, those with toroidal to poloidal mode number ratios of n/m = 1/5, 1/4 and 1/3 and thus the dominant periodicity of the magnetic field structure). A few islands with toroidal mode number n = 2 were also detected in the database cases, e.g., those with the mode structure 2/5 and 2/7. While the 2/7 islands, with a large m value, had exceedingly small widths and therefore were unsuited to be determined even with a fine resolution, the 2/5 islands, though with larger widths, were (at least) as small as the 1/5 islands. As will be described later, the 1/5-island widths already involved quite large relative errors and so the 2/5 islands were not used in the database.

Choice of the aforementioned variables for analysis was based on their importance in control purposes for which the FP equations are to be used. It may be commented here that the magnetic field

strength has only a weak toroidal dependence in WEGA, so the toroidal mirror field effect is negligible for both the ideal coil system and the considered misaligned coil system.

All variables listed above have been locally determined, at a particular point of the configuration, and are called scalar variables. Only the rotational transform ι was analysed as a profile quantity with a dependence on r_{eff}, which is an effective minor radius for labelling flux surfaces. Its definition, using the cross sectional areas of the flux surfaces, is the same as that mentioned in [3]. Mathematically, the ι-dependency was considered to be with r_{eff}^2. This is because profile variables in magnetic configurations are strictly functions of magnetic flux enclosed by the flux surfaces and r_{eff}^2, compared to r_{eff}, is the main radial dependence of the flux.

Definitions of r_{is} and w_{is} were the same as those in [3], namely, from the detection of the inner and the outer separatrices of a magnetic island chain. If the r_{eff} values corresponding to these separatrices are $r_{eff}^{(innsep)}$ and $r_{eff}^{(outsep)}$ then

$$r_{is} = \frac{r_{eff}^{(outsep)} + r_{eff}^{(innsep)}}{2} \qquad \text{and} \qquad w_{is} = r_{eff}^{(outsep)} - r_{eff}^{(innsep)} \tag{4}$$

Therefore, an accurate detection of the inner and the outer separatrices is a very important criterion for calculating the island parameters. Very often, however, the outer separatrix is not clearly formed due mainly to stochastic regions surrounding, e.g., islands of large size which usually also means large values of ι. This uncertainty makes especially the island width somewhat inaccurate in the database which is also reflected in the error of the FP approximations.

A field line tracing code [9], whose inputs were the currents in the TF, HF and VF coils, was used to generate the vacuum magnetic field. The misalignment of 4 mm between the centres of the TF and the HF coils was included in the numerical model to reproduce the measured error field effects.

Since magnetic configurations are essentially invariant under a global scaling of the coil currents, except for the magnetic field whose strength varies explicitly with the current in the relevant coil, the predictor variables (components of $\vec{x}$) for the FP models were the CCR $i_{HF} = I_{HF}/I_{TF}$ and $i_{VF} = I_{VF}/I_{TF}$, where i and I denote CCR and the coil currents, respectively. The TF coil is maintained at a constant current in the experiment, and so it was in the dataset (at $I_{TF} = 2$ kA) and formed the normalisation parameter, so the magnetic field strength should be duly scaled according to the experimental value of I_{TF}. For CCR, the ranges $0.5 \leq i_{HF} \leq 2.7$, and $-0.05 \leq i_{VF} \leq 0.05$ were used in generating the dataset of 250 configurations for our analysis.

Figure 2 shows the configuration space for CCR. Three aspects of the scatter plot are worth noting. First, the data points were not generated in the pseudo-random way as is often done for statistical analyses, but systematically in a 2-D grid defined by i_{HF} and i_{VF}. This is justified due to the low dimensionality of the input space. Assuming that a polynomial of order p is necessary to represent an output variable, the criterion for the minimum number of points in the configuration space for the regression is p^2 in our case, and with the order of 100 points in the space the criterion is safely satisfied. Second, there is a void in the bottom-left corner. This region, at small i_{HF}, corresponds to ultra-low values of rotational transform ι so that flux surfaces are not well-formed due to insufficient twist in the magnetic field lines. As seen from figure 3, ι_{ax} increases with i_{VF}, so the void does not continue upwards. There are a few additional voids for $i_{HF} = 2.1$ and 2.3. For these configurations the code failed in the method of an
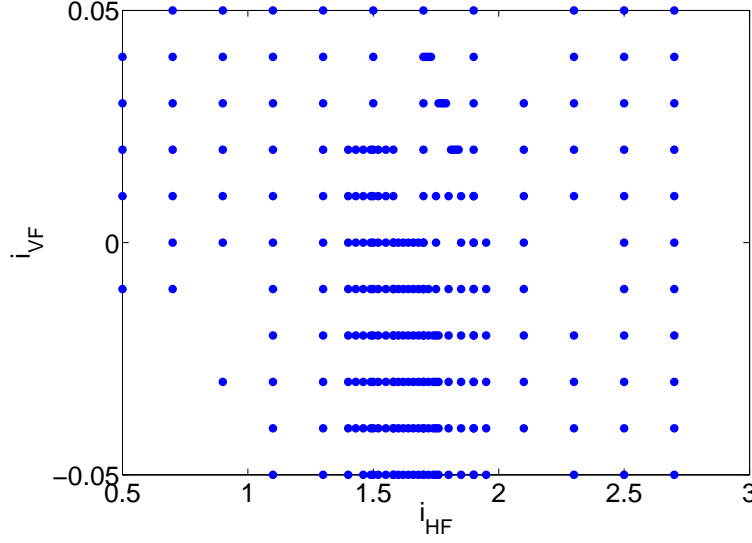
Figure 2: WEGA configuration space in the coil current ratios.

automatic determination of the separatrix. Small shifts of these points in the grid for these cases, e.g., by altering i_{HF}, may produce successful runs. However, this was not considered necessary because, even without those grid points, the space is still reasonably uniformly covered. Third, a clustering of points is seen in the region with $1.4 \leq i_{HF} \leq 2.0$. These points were generated in order to get a good coverage of the part of the configuration space containing the magnetic islands of interest, namely those with the modes 1/3, 1/4 and 1/5.

Table 1 lists the summary statistics for the configuration parameters in the database, that includes the mean, the standard deviation or spread σ and the extreme values. The CCR were varied so that the physical state parameters at least encompass all the experimental scenarios. The values of ℓ, for example, cover with values between 0.017 and 0.837 the experimental range for the low field scenario; in case of the high field operation the experimental range is much smaller (typically $0.20 \leq \ell < 0.50$). The a_{eff} values also span a wide range, from very small configurations ($a_{eff} < 2$ cm) up to very large ones ($a_{eff} > 15$ cm). The former correspond to high ℓ cases which imposes an upper limit on the ℓ values since these configurations are too small to be of experimental use. The larger a_{eff} values in the data safely exceed the present experimental limit of 11 cm. For the magnetic island locations, cases with islands too close to the magnetic axis were excluded from the dataset as their parameters are anyway difficult to be determined, even in the experiment.

One of the configurations in the database, corresponding to $(i_{HF}, i_{VF}) = (1.76, -0.04)$, is shown in a Poincaré plot in figure 4, at $\phi = 0$. This clearly displays the symmetry-breaking, and not the natural 5-fold periodic, island chains (of modes 1/3 and 1/4). The corresponding rotational transform profile in figure 5 shows the considerable positive shear in this configuration (defined as the gradient of the ℓ-profile in the r_{eff} space). For WEGA it has been observed that, as i_{VF} is moved from negative through zero to positive values (at constant i_{HF}), the shear keeps reducing so that for the upper regions of the configuration space in our database (figure 2) the ℓ-profile is almost flat with only a small upward trend close to the boundary.
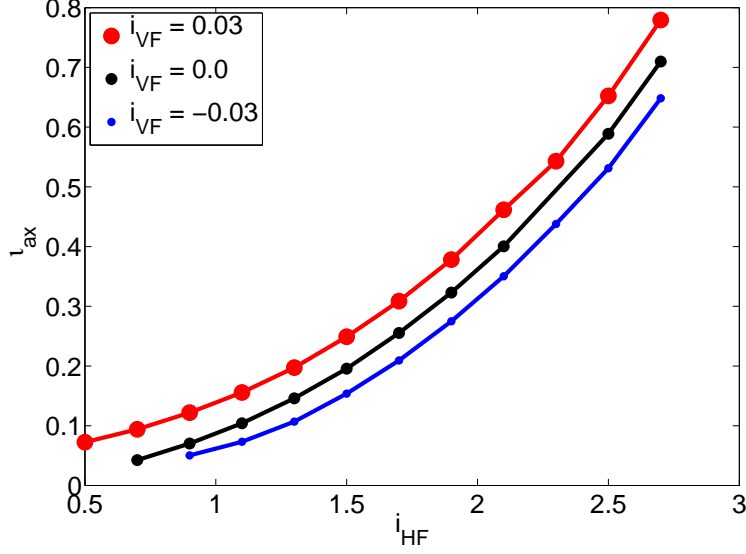
Figure 3: Variation of ι_{ax} with i_{HF} for no vertical field ($i_{VF} = 0$) and two equal and opposite vertical fields ($i_{VF} = \pm 0.03$).
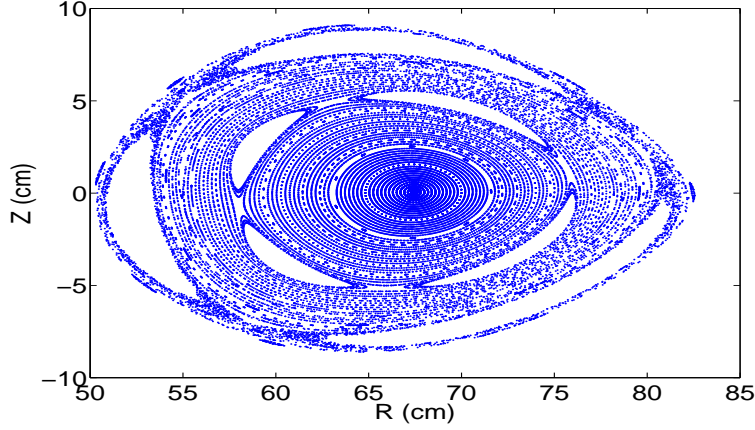

Figure 4: Poincaré plot for case number 202 ($i_{HF} = 1.76$, $i_{VF} =$ -0.04) in the WEGA vacuum database. Islands with 1/3 and 1/4 modes are clearly visible.

4. STATISTICAL ANALYSIS RESULTS

As explained in section 2, the purpose of statistical analysis on a dataset is to develop global (i.e., over the entire configuration space defined in the dataset) representations of physical quantities in terms of the predictors, and a study of the error statistics leads to a decision on the best fit. These will now be described for the physical parameters of WEGA. From the database, 175 configurations were used for setting up the FP models, making sure the parameter values covered at least all relevant experimental cases, and another 75 to test the quality of fit. We would like to point out that the external coil currents were assumed to be accurately measurable so the simulated coil currents were not perturbed with measurement errors. This was reasonable, since coil current measurements usually involve only very small levels of uncertainties.

Prior to setting up the statistical models, an exploratory PCT was performed on the CCR in the database to test whether they, having been non-randomly generated in a grid, are uncorrelated. Eigen-analysing the 2×2 correlation matrix generated out of the CCR data, it was found that (a) the first of the
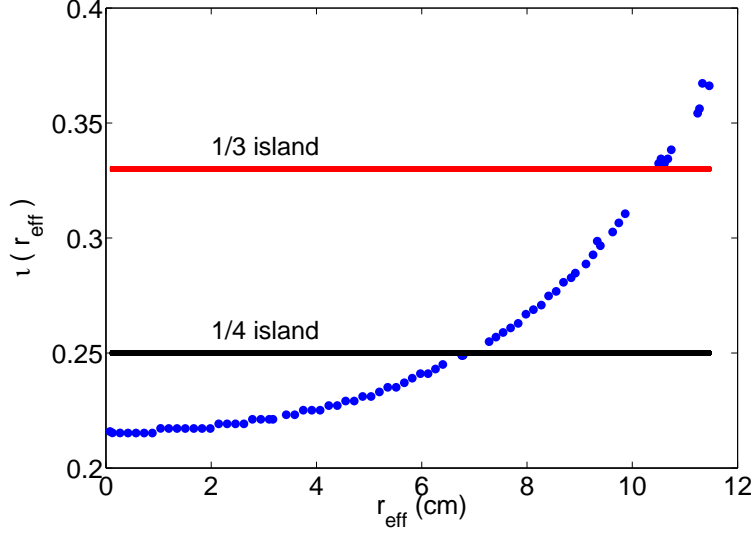
Figure 5: The ι profile for case number 202 in the WEGA vacuum database. Discontinuities are at the locations of islands. The configuration has a large, positive shear.

two principal components explained only 58% of the total variance of i_{HF} and i_{VF}, (b) the ratio of the maximum eigenvalue of the correlation matrix to the minimum, called the condition number, was 1.37, i.e., of the order of unity, signifying that the CCR data were well-conditioned, and (c) the magnitude of the correlation coefficient of i_{HF} and i_{VF} was 0.15. Thus, the raw "measurements" i_{HF} and i_{VF} were indeed only weakly correlated in the database and so were fit to be used as the independent variables in the regression model.

a) Scalar parameter recovery

The regressions were done and tested with a quadratic (2-FP), a cubic (3-FP), a 4th order (4-FP) and a 5th order (5-FP) polynomial in (i_{HF}, i_{VF}) including interaction terms, that involved 6, 10, 15 and 21 model coefficients, respectively. In [3] and [4] expressions were given to combine the predictors in quadratic and cubic. When the fourth and the fifth order terms, including interactions, are added, the expressions expand into a 4-FP and a 5-FP polynomial of the forms

$$y = \sum_{i=0}^{n_{in}} \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} a_{ijkl} \, x_i x_j x_k x_l \tag{5}$$

and

$$y = \sum_{i=0}^{n_{in}} \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} \sum_{p=0}^{l} a_{ijklp} \, x_i x_j x_k x_l x_p \tag{6}$$

respectively. In the two equations above, n_{in} is the number of (uncorrelated) predictors.

Table 2: Recovery statistics of scalar parameters using linear regression									
Parameter	spread	RMS error from models				100 (RMS error)/(spread)			
		2-FP	3-FP	4-FP	5-FP	2-FP	3-FP	4-FP	5-FP
ι_{ax}	0.215	0.0060	0.0035	0.0020	0.0018	2.79	1.63	0.93	0.84
ι_b	0.218	0.0160	0.0126	0.0113	0.0108	7.34	5.78	5.18	4.95
a_{eff}	4.40 cm	1.70 cm	1.33 cm	1.0 cm	0.93 cm	38.29	30.00	23.50	21.40
R_{ax}	3.95 cm	1.66 cm	1.00 cm	0.59 cm	0.30 cm	42.00	25.30	14.94	7.59
B_{ax}	0.018 T	0.0073 T	0.0041 T	0.0028 T	0.0019 T	40.56	22.78	15.56	10.56
$r_{is}^{(1/3)}$	2.44 cm	—	0.48 cm	0.37 cm	0.20 cm	—	19.85	15.30	8.12
$w_{is}^{(1/3)}$	1.15 cm	—	0.80 cm	0.43 cm	0.35 cm	—	69.23	37.79	30.00
$r_{is}^{(1/4)}$	2.06 cm	—	0.45 cm	0.18 cm	0.11 cm	—	21.74	8.74	5.10
$w_{is}^{(1/4)}$	0.45 cm	—	0.16 cm	0.12 cm	0.07 cm	—	36.19	26.03	14.50
$r_{is}^{(1/5)}$	2.57 cm	—	0.12 cm	0.08 cm	0.08 cm	—	4.28	2.61	2.49
$w_{is}^{(1/5)}$	0.41 cm	—	0.10 cm	0.09 cm	0.09 cm	—	24.39	21.95	19.51

In our previous studies [3, 4] on W7-X configurations, a 3-FP model was always found to be necessary and sufficient. For WEGA a significant improvement in the regression accuracy was observed, for all the configuration parameters regressed, when a 4-FP model was used. The error statistics for all regressed parameters are tabulated in Table 2, where the spread σ, the root-mean-square (rms) error ϵ_{rms}, the R^2-measure of fit (which is the fraction of the total variance of the regressed variable explained by the model) and the percentage spread error ϵ_{perc} (defined as ϵ_{rms} normalised to the database spread σ) are recorded. The last-mentioned statistic is given by

$$\epsilon_{perc} = 100 \times (\epsilon_{rms}/\sigma) = 100 \times \sqrt{1 - R^2} \qquad (7)$$

where the R^2-measure of fit is the fraction of the total variance of the regressed variable explained by the model.

The central iota ι_{ax} shows a progressive improvement with the size of the model, up to 4-FP when its estimation is accurate to within ± 0.002 corresponding to $\epsilon_{perc} = 0.93\%$ and an R^2 statistic of 0.9999. The slow decrease of the error level seen for the 5-FP may indicate that the accuracy of the numerical procedure to determine the iota-value is in reach.

The boundary parameters, due to an inherent uncertainty of locating the boundary of the configuration, could not be very accurately generated in the data. This inaccuracy is more pronounced if the configuration is bounded by a separatrix. This feature inevitably shows up in the recovery and the two parameters in question, the ι_b and a_{eff}, were less accurately estimated. For ι_b the recovery was correct to within ± 0.01 corresponding to $\epsilon_{perc} = 5.7\%$ and an R^2 measure of fit of 0.9967 with 3-FP model, and
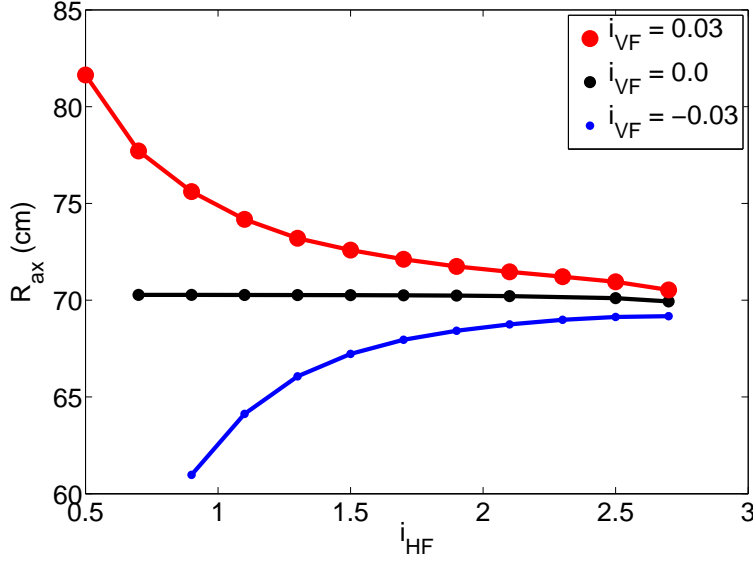
Figure 6: Variation of R_{ax} with i_{HF} at three fixed values of i_{VF}.

did not improve significantly with the extra non-linear terms of the higher order models. The sufficiency of 3-FP for the regression of ι_b is also manifested by the fact that 7 of the 15 regression coefficients of the 4-FP model were found to be statistically insignificant at the 5% level, when tested against the null hypothesis that the parameter values are zero. This means that the significance probabilities for these coefficients were greater than 0.05 and so the null hypothesis could not be rejected. For a_{eff}, the recovery errors significantly reduced up to 4-FP, when the estimation of the parameter is found to be correct to ± 1 cm. Hence, with increasing model size the recovery errors of both variables approach the error level of the dataset due to the imperfect estimation of these quantities.

From the table it is also clear that for R_{ax}, B_{ax}, and for the parameters of at least two island modes there was a further significant enhancement in the quality of regression with a 5-FP model. The necessity of higher order regression terms in the models suggests either strong non-linear dependencies of the physical parameters of WEGA configuration on the coil currents, or that the polynomial function may not be the best choice for fitting. The latter is supported by the dependency that R_{ax} shows in figure 6. We clearly see that R_{ax} behaves approximately like $1/i_{HF}$. Therefore, a polynomial model in i_{HF} that represents this variation is expected to need significant higher order terms for a good approximation. However, the mere use of $1/i_{HF}$ as the regressor, instead of i_{HF}, will not achieve success since the singularity may not be at $i_{HF} = 0$ and additionally it changes sign with i_{VF}. This suggests the use of rational functions to model this behaviour.

Before continuing further, some comments on the island parameters should be made here. First, we did not bother to use a 2-FP model for r_{is} and w_{is}, as we presumed that their regression will demand many more non-linear terms. The 3-FP error statistics (Table 2) justify this. Second, for the 1/4 and 1/5 modes a regression of r_{is}^2 expectedly produced significantly better results – and these are the ones quoted in the table. The expectation is based on the observations that the profile of iota behaves like r_{eff}^2 around the magnetic axis (figure 5) and that the CCR-dependence of ι_{ax} may be approximated by a linear function around the value where an island appears. Figure 7 shows the island location for the three modes as functions of i_{HF} for i_{VF} = -0.04.
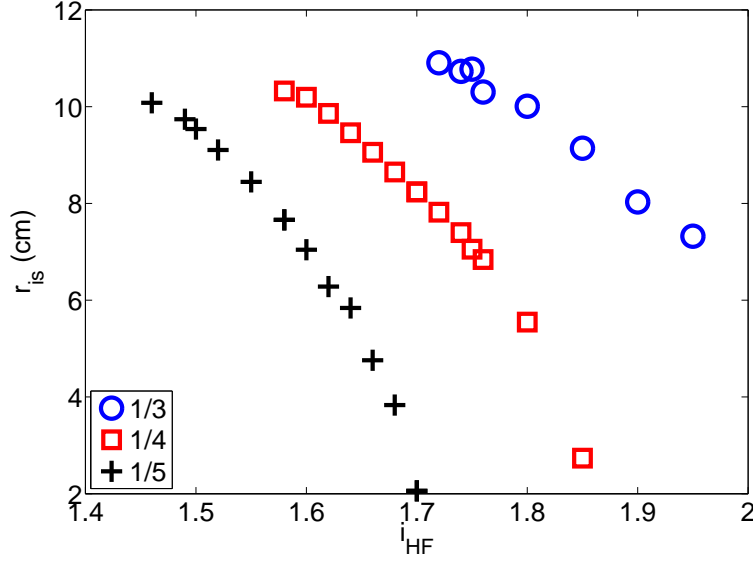
Figure 7: Variation of island locations with i_{HF} (or t) at i_{VF} = -0.04. Consequences of positive shear is visible.

For all the parameters whose regression showed progressively better fits even up to 5-FP, we decided to test the feasibility of non-linear regression using rational functions. The purposes were (i) to study the usefulness of these functions from the viewpoint of accuracy as compared to, e.g., the 5-FP model, and (ii) to test the size of each regression model compared to the 21-coefficient 5-FP. Non-linear regressions involve iterative convergence of the error function to its minimum in the hyperspace of the model coefficients, and the Levenberg-Marquardt scheme was used for that. The results were very encouraging, and the rational functions were finally settled with as the best fit.

Denoting i_{HF} and i_{VF} by x_1 and x_2 respectively, for convenience, the best fit rational functions for the relevant configuration parameters were the following :–

$$R_{ax} = \frac{a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1^2 + a_4 x_2^2}{a_5 + a_6 x_1 + a_7 x_2 + a_8 x_1^2 + a_9 x_2^2 + a_{10} x_1 x_2 + a_{11} x_1^3 + a_{12} x_2^3} \tag{8}$$

with 13 coefficients, ϵ_{rms} = 1.1 mm, R^2 = 0.9992, ϵ_{perc} = 2.84;

$$B_{ax} = \frac{b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_2^2}{b_5 + b_6 x_1 + b_7 x_2 + b_8 x_1^2 + b_9 x_2^2 + b_{10} x_1^3} \tag{9}$$

with 11 coefficients, ϵ_{rms} = 0.0007 T, R^2 = 0.9988, ϵ_{perc} = 3.48;

$$r_{is}^{(1/3)} = \frac{c_0 + c_1 x_1 + c_2 x_2 + c_3 x_1^2 + c_4 x_2^2}{c_5 + c_6 x_1 + c_7 x_2 + c_8 x_1^2 + c_9 x_2^2} \tag{10}$$

with 10 coefficients, ϵ_{rms} = 1.8 mm, R^2 = 0.9932, ϵ_{perc} = 8.27;

$$w_{is}^{(1/3)} = \frac{d_0 + d_1 x_1 + d_2 x_2 + d_3 x_1^2 + d_4 x_2^2 + d_5 x_1^3 + d_6 x_2^3}{d_7 + d_8 x_1 + d_9 x_2 + d_{10} x_1^2 + d_{11} x_2^2 + d_{12} x_2^3} \tag{11}$$

with 13 coefficients, ϵ_{rms} = 2.1 mm, R^2 = 0.9515, ϵ_{perc} = 22;

$$r_{is}^{(1/4)} = \frac{e_0 + e_1 x_1 + e_2 x_2 + e_3 x_1^2 + e_4 x_2^2 + e_5 x_1 x_2}{e_6 + e_7 x_1 + e_8 x_2 + e_9 x_1^2 + e_{10} x_2^2} \tag{12}$$

with 11 coefficients, $\epsilon_{rms} = 0.96$ mm, $R^2 = 0.9973$, $\epsilon_{perc} = 5.18$;

$$w_{is}^{(1/4)} = \frac{g_0 + g_1 x_1 + g_2 x_2 + g_3 x_1^2 + g_4 x_2^2 + g_5 x_1 x_2}{g_6 + g_7 x_1 + g_8 x_2 + g_9 x_1^2 + g_{10} x_2^2} \qquad (13)$$

with 11 coefficients, $\epsilon_{rms} = 0.76$ mm, $R^2 = 0.9782$, $\epsilon_{perc} = 14.7$.

The R^2-statistic in equations (9) - (14) were calculated from the basic definition in equation (8).

For R_{ax} and B_{ax} there was a remarkable improvement in accuracy over a 5-FP model, while for the island parameters the 5-FP results were reproduced, by the use of rational functions with 8-10 coefficients less. This result suggests that the rational functions are indeed better approximators of the dependencies for R_{ax} and B_{ax}. For the island parameters the uncertainties within the database set the level of the achievable accuracy; this is inferred from the fact that the error levels obtained from the 5-FP model and the rational functions are more or less the same, except for $w_{is}^{(1/3)}$. For this particular parameter the accuracy was indeed improved from 3.5 mm with a 5-FP model to 2.1 mm with the rational function. Generally, the errors in the data leading to a poorer recovery of the island widths can be explained as follows. For the modes with high ι, the region outside the island chain is usually stochastic and as such the island may not be properly formed, leading to a possible misjudgement on the location of the outer separatrix. This was reflected in the recovery of $w_{is}^{(1/3)}$. On the other hand, the islands with moderate to high poloidal mode numbers m (which generally also implies moderate to low ι for fixed n) do not have very large widths. In fact, the mean width for the 1/5 islands (the ones with the largest m) in our database was only 6 mm. However, this smallness may also result in large relative errors in their determination, as was observed for the 1/5 island widths. The 1/4 islands, being in-between, were less affected from both the aforementioned problems and so their width was determined with better relative accuracy which was passed on to the accuracy of the recovery.

The superiority of the rational functions was also proved in a comparison with a different approach, namely, regressing the inverse of the original response variables, especially where a $1/x$-dependence was observed. For example, the regression of $1/R_{ax}$, instead of R_{ax}, improved the rms error from 3 mm to 2.2 mm using 5-FP, but the result of the rational function FP is still better by a factor of 2.

The use of higher order polynomials or rational functions improves the accuracy of the recovered response variables, however, they are less stiff compared to the standard 2-FP and extrapolation bears the risk of getting unreasonable values. This is especially important for rational functions since they will very likely have singularities close to the parameter region covered by the dataset if their non-linear behaviour is important to describe the trend. For control purposes, the allowed parameter range for the FP-usage has to be restricted to the safe parameter regions, on which the training dataset was well approximated.

b) Profile parameter recovery

As already listed, the only relevant profile quantity for WEGA is ι which, in the context of the FP equation, was considered to be a function of CCR and r_{eff}^2, as already explained. The ι-profile was parametrized with a radial polynomial of the form

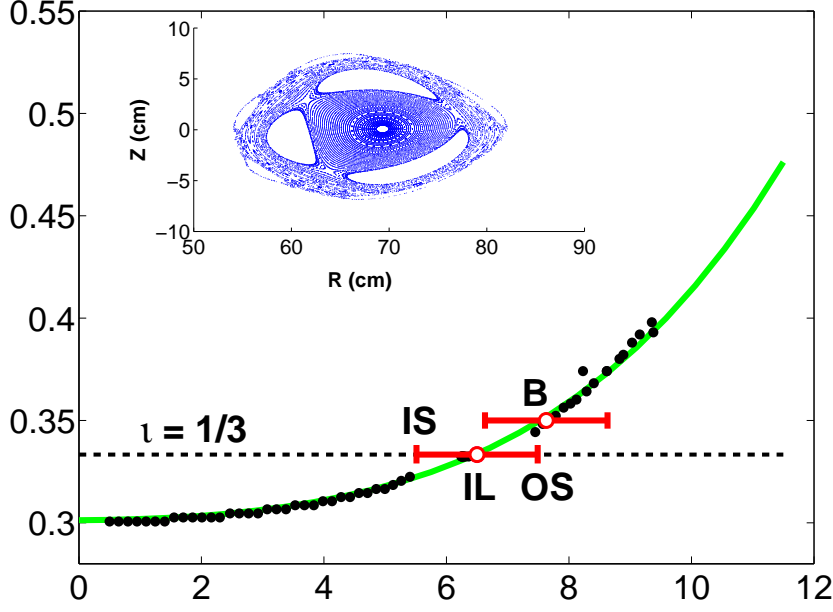$$\iota(i_{HF}, i_{VF}, r_{eff}^2) = p_0 + p_1 r_{eff}^2 + p_2 (r_{eff}^2)^2 \qquad (14)$$

Figure 8: FP-reconstructed ι-profile (blue) for a configuration generated with $(i_{HF}, i_{VF}) = (1.9488, -0.0243)$, and the corresponding data points (black filled circles) from code calculations. Also shown are the approximate limit B (red filled circle) of good flux surfaces from predicted a_{eff} including its uncertainty and the location IL (magenta filled circle) of the 1/3 mode. IS and OS denote the inner and the outer separatrix, respectively. Inset: Poincaré plot for this configuration for comparison.

where CCR (i_{HF}, i_{VF}) are contained within the coefficients p_i. When CCR are combined in a "mixed" cubic form, the uncertainty of ι over the entire profile was ± 0.0025; when the combination was in a "mixed" 4th order form, the uncertainty was ± 0.0015. So the latter was chosen as the best fit, consistent with the scalar parameter results. Higher powers of r_{eff}^2 did not improve the regression at all.

Some features of the radial regression function include the following:

a) It is a smooth function and so unable to represent the discontinuity in the profile due to the effect of islands.

b) Being a quadratic function in r_{eff}^2, which is rather stiff, it mathematically allows extrapolations beyond the LCFS of a given configuration. It will return a value of ι for any r_{eff} whether flux surfaces exist for that r_{eff} in the real configuration or not.

To avoid the misinterpretation b) of the regression function it is necessary to combine the regression of ι-profile with that of the island location and/or a_{eff} in order to restrict the valid range of r_{eff} for the regression functions.

Figure 8 is a demonstration of the above. The blue curve is the reconstructed ι-profile for a configuration in the test-data (i.e., one which is not in the training dataset) generated by $(i_{HF}, i_{VF}) = (1.95, -0.02)$. The curve smoothly continues indefinitely in r_{eff} with realistic values of ι. However, the predicted a_{eff} (point B) is only 7.62 cm (with an uncertainty of $\epsilon_{rms} = \pm 1$ cm) that denotes the FP-predicted extent along the profile to which good flux surfaces should exist in the configuration. A 1/3 island is predicted from the FP-$r_{is}^{(1/3)}$ to be at $r_{eff} = 6.50 \pm 0.18$ cm. The island location is indicated by IL in the figure. This estimate is consistent with the appeerence of the appropriate rational surface in the reconstructed ι-profile as seen from the figure. The island width $w_{is}^{(1/3)}$ is predicted to be 1.98

cm with $\epsilon_{rms} = \pm 0.21$ cm. The points IS and OS, drawn at distances $w_{is}^{(1/3)}/2$ and $-w_{is}^{(1/3)}/2$ from IL, denote the inner and the outer separatrices, respectively. From the predicted a_{eff}, $r_{is}^{(1/3)}$ and the island width, it appears that the configuration has a few good flux surfaces beyond the island chain.

The inset to figure 8 is the Poincaré plot for the configuration to compare with the above predictions. It shows the configuration to be separatrix-bound by a 2/5 island mode while the 1/3 island chain is internal. The filled black circles are the data for the ι-profile from code calculations. The indicated island size seems to be consistent with the gap in the recalculated ι-profile for which the line of starting points passed through the x-point. Up to the inner separatrix of the 2/5 island, a_{eff} is found to be ~ 9 cm, so the FP-predicted value of 7.62 cm is somewhat underestimating it. The discrepancy is possibly due to the fact that the region of configuration space to which this configuration belongs is one where a_{eff} is found to have steep gradients in its variation with i_{HF} and where the ι-profile has a considerable shear at larger r_{eff}. Since the FP model for a_{eff} is a smooth (polynomial) function of CCR, it may have had problems in a more accurate prediction. This is in addition to the already stated reason of boundary-related parameters being somewhat erroneous in the data itself due to the uncertainty in locating the boundary.

Reasonable consistency and agreement on predicting the locations of the rational surfaces using the models for the ι-profile and r_{is} have also been observed. A 1/3 island case has been already described above. For a 1/4 island mode, with $(i_{HF}, i_{VF}) = (1.75, -0.03)$, we obtained $\iota(r_{eff}) \sim 0.25$ at $r_{eff} \sim 6.05$ cm. For the same CCR, $r_{is}^{(1/4)} = 6.04 \pm 0.096$ cm. To get the location of a 1/5 mode, a configuration generated by $(i_{HF}, i_{VF}) = (1.56, -0.02)$ was used. The ι-profile model predicted the mode rational surface to be at $r_{eff} \sim 7.13$ cm, while the predicted $r_{is}^{(1/5)}$ was 7.05 ± 0.08 cm.

5. CONCLUSIONS

The FP on WEGA, though performed on a small device with a small number of independently variable predictors, has shown the inherent non-linearities in the dependencies of the physical state of the magnetic configurations on the external coil currents. We successfully modelled the parameters of the low-order rational islands in the configurations generated by the symmetry-breaking error fields due to the misalignment of the centres of the toroidal and helical field coil systems. The best-fit statistical models needed to be either 4th order polynomials with linear regression, or rational functions with non-linear regression. The fact that the latter, for some of the physical parameters, improved upon the 4th and even the 5th order polynomial regressions demonstrated the strong non-linearities in the dependency. Even then, all the regression models were of modest size due to the small number of predictors, with a 4-FP having 15 estimated regression coefficients being the largest. Our results, therefore, were encouraging. However, extrapolation of the high order polynomials and rational functions has to be avoided wespecially for use in a control system. Modelling the magnetic islands was an important part of this study, and this was quite successful except, to some extent, for the challenging issue of the island width whose "measurement" (database) itself can be erroneous due to the problems linked with the detection of the outer separatrix. Since coil currents were the only measurements involved, noise in the predictors was neglected as these are usually very small. The validity of an important assumption of the theory of statistical regressions, namely, that of the predictor variables being measured without error, was thus

maintained. The FP functions are now in the process of being implemented in software to run with the control system.

REFERENCES:

[1] Otte M., Lingertat J., 29th EPS Conference on Plasma Phys. and Controlled Fusion, Montreux, ECA, Vol.26B, P-5.036 (2002)

[2] Podoba Y.Y., et al, *Physical Review Letters* **98** (2007) 255003

[3] Sengupta A., et al, *Nuclear Fusion* **44** (2004) 1176

[4] Sengupta A., et al, *Plasma Phys. Contr. Fusion* **49** (2007) 649

[5] Braams B.J., Jilge W., Lackner K., *Nuclear Fusion* **26** (1986) 699

[6] van Milligen B.Ph. and Lopes Cardozo N.J. *Comput. Phys. Commun.* **66** (1991) 243

[7] McCarthy P.J. and Morabito F.C. *Int. J. Appl. Electromag. Mech.* **8** (1997) 343

[8] Mardia K.V., Kent J.T., Bibby J.M., Multivariate Analysis, Academic Press, New York (1979)

[9] Werner A., IPP-Greifswald, Germany (2007): private communications.