

A Hierarchy of Empirical Models of Plasma Profiles and Transport

K Imre¹, K S Riedel¹, B Schunke.

JET Joint Undertaking, Abingdon, Oxfordshire, OX14 3EA, UK.

¹ New York University, 251 Mercer St., New York, NY 10012-1185, USA.

Preprint of a paper accepted for publication in
Physics of Plasmas, May 1995

December 1994

"This document is intended for publication in the open literature. It is made available on the understanding that it may not be further circulated and extracts may not be published prior to publication of the original, without the consent of the Publications Officer, JET Joint Undertaking, Abingdon, Oxon, OX14 3EA, UK".

"Enquiries about Copyright and reproduction should be addressed to the Publications Officer, JET Joint Undertaking, Abingdon, Oxon, OX14 3EA".

ABSTRACT

Two families of statistical models of increasing statistical complexity are presented which generalize global confinement expressions to plasma profiles and local transport coefficients. The temperature or diffusivity is parameterized as a function of the normalized flux radius, $\bar{\psi}$, and the engineering variables $\mathbf{u} = (I_p, B_t, \bar{n}, q_{95})^\dagger$. The log-additive temperature model assumes that $\ln[T(\bar{\psi}, \mathbf{u})] = f_o(\bar{\psi}) + f_I(\bar{\psi})\ln[I_p] + f_B(\bar{\psi})\ln[B_t] + f_n(\bar{\psi})\ln[\bar{n}] + f_q \ln[q_{95}]$. The unknown $f_i(\bar{\psi})$ are estimated using smoothing splines. The Rice selection criterion is used to determine which terms in the log-linear model to include. A 43 profile JET Ohmic data set from the Joint European Torus [P.H. Rebut, et al., Nuclear Fusion **25** 1011, (1985)] is analyzed and its shape dependencies are described. Our best fit has an average error of 152 eV which is 10.5 % percent of the typical line average temperature. The average error is less than the estimated measurement error bars. Our second class of models is log-additive diffusivity models where $\ln[\chi(\bar{\psi}, \mathbf{u})] = g_o(\bar{\psi}) + g_I(\bar{\psi})\ln[I_p] + g_B(\bar{\psi})\ln[B_t] + g_n(\bar{\psi})\ln[\bar{n}]$. These log-additive diffusivity models are useful when the diffusivity is varied smoothly with the plasma parameters. We describe a penalized nonlinear regression technique to estimate the $g_i(\bar{\psi})$. The physics implications of the two classes of models, additive log-temperature models and additive log-diffusivity models, are different. The additive log-diffusivity models adjust the temperature profile shape as the radial distribution of sinks and sources. In contrast, the additive log-temperature model predicts that the temperature profile depends only on the global parameters and not on the radial heat deposition.

PACS NUMBERS:02.52.55Fa, 52.55Pi, 52.65+z

I. INTRODUCTION

Global confinement expressions have proven useful in understanding and predicting plasma performance¹⁻⁷. These confinement expressions are straightforward to analyze statistically, but do not address the radial variation of the plasma profiles and plasma transport coefficients. In this article, we describe two families of empirical models which generalize global scaling expressions to profiles and diffusivities.

We define the m vector, \mathbf{u} , to be a vector of global engineering variables. Typically, the components of \mathbf{u} are the logarithms of the edge safety factor, q_{95} , the plasma current, I_p (in MA), the toroidal magnetic field, B_t (in Tesla), the line average density, \bar{n}_e (in $10^{19}/\text{m}^3$), the absorbed power, P (in MW), the effective ion charge, Z_{eff} , the isotope mixture, M , the plasma elongation, κ , and the major and minor radii. Other engineering variables can include the divertor configuration, the wall type, and the type of heating. In practice, we usually work with the logarithms of the engineering variables, and normalize the variables about their mean values in the data set. In this notation, the standard power law for the energy confinement time, τ_E , is

$\tau_E = c_0 I_p^{\beta_1} B_t^{\beta_2} \bar{n}^{\beta_3} \dots$, where β_ℓ are the scaling exponents. The power law can be rewritten as a log-linear expression

$$\ln \tau = \beta_0 + \beta_1 \ln[I_p] + \beta_n \ln[\bar{n}] + \dots, \quad (1)$$

In this form, the resulting scaling expression can be analyzed using linear regression. Ordinary least squares analysis makes a number of implicit assumptions which are described in Refs. 1 and 3.

We consider the plasma temperature as a function of the normalized radial flux variable, $\bar{\psi}$, and the plasma control variables, \mathbf{u} . A convenient flux variable normalization is that the toroidal flux through a given radius, $\bar{\psi}$, is equal to $\bar{\psi}^2$ times the total flux. In this section, we neglect the random errors associated with the data and concentrate on the empirical model. By writing $T(\bar{\psi}, \mathbf{u})$, we are implying that the temperature is an unknown function of $m+1$ variables. Attempting to estimate an arbitrary $m+1$ dimensional function from the measured tokamak data is a very ill-conditioned problem. Therefore, we restrict the class of models which we examine to a more limited class.

Experimentalists have often observed that the Ohmic temperature profile shape varies very little as the plasma control variables vary. This ‘‘profile resilience’’ motivates us to define ‘‘profile resilient’’ models:

$$\ln[T(\bar{\psi}, \mathbf{u})] = f_0(\bar{\psi}) + H(\mathbf{u}) \quad (2)$$

$H(\mathbf{u})$ is independent of $\bar{\psi}$ and changes the magnitude but not the shape of the temperature profile. In the log-linear case, $H(\mathbf{u}) = c_I \ln[I_p] + c_B \ln[B_t] + c_n \ln[\bar{n}] + \dots$. We determine $f_0(\bar{\psi})$ and $H(\mathbf{u})$ by fitting $f_0(\bar{\psi})$ with smoothing splines and using linear regression.

A more general model is to let the shape depend on q_{95} :

$$\ln[T(\bar{\psi}, \mathbf{u})] = f_0(\bar{\psi}) + f_q(\bar{\psi})q_{95} + H(\mathbf{u}) \quad (3)$$

where both unknown radial functions, $f_0(\bar{\psi})$ and $f_q(\bar{\psi})$, are fit with smoothing splines. Tang's transport model^{8,9}, based on profile consistency, is a special case of model of Eq.(3). Tang's model requires the log-temperature profile shape to be quadratic: ($f_0(\bar{\psi}) \equiv c_0 \bar{\psi}^2$ and $f_q(\bar{\psi}) \equiv c_q \bar{\psi}^2$), and derives $H(\mathbf{u})$ from theoretical consideration.

More generally, we have the additive spline model of Refs. 10-12:

$$\ln[T(\bar{\psi}, \mathbf{u})] = f_0(\bar{\psi}) + H(\mathbf{u}) + \sum_{\ell=1}^L f_\ell(\bar{\psi})h_\ell(\mathbf{u}) \quad (4)$$

In Eq. (4), we have separated $f_0(\bar{\psi})$ and $H(\mathbf{u})$ from the other terms to stress that these terms are the “profile consistent” terms. If desired, additional cross-terms may be added to Eq.(4).

In Section II, we describe our fitting procedure for estimating the free parameters in the log-additive model of Eq. (4). In Section III, we apply our method to the Ohmic data from the Joint European Torus¹³ (JET) with $\mathbf{h}(\mathbf{u}) = (\ln[I_p], \ln[B_t], \ln[\bar{n}], \ln[q_{95}])$. The resulting model can be rewritten as

$$T(\bar{\psi}) = \mu_0(\bar{\psi}) I_p^{f_1(\bar{\psi})} B_t^{f_2(\bar{\psi})} \bar{n}^{f_3(\bar{\psi})} q_{95}^{f_4(\bar{\psi})}, \quad (5)$$

where $\mu_0(\bar{\psi}) = \exp(f_0(\bar{\psi}))$. In place of q_{95} in Eq. (5), we can use the geometric part of the safety factor: $\hat{q}_{\text{geo}} \equiv q_{95} I_p / B_t$. We also find that a simpler model with “ f_1, f_2 and f_3 =constant” fits the data to reasonable precision.

In Sections IV and V, we introduce a second family of models for the log-diffusivity. We summarize our results in Sec. VI. The Appendix describes our model selection criteria; i.e. how we use the data to determine which log-linear model is most appropriate.

II. PROFILE ESTIMATION AND MODEL SELECTION

To estimate the unknown functions, $f_\ell(\bar{\psi})$ in the additive log-temperature models, we expand each of the functions in B-splines: $f_\ell(\bar{\psi}) = \sum_{k=1}^K \alpha_{\ell k} B_k(\bar{\psi})$, where the $B_k(\bar{\psi})$ are the cubic B-spline functions. The $\alpha_{\ell k}$ are free parameters which need to be estimated. In Refs. 10-12, we describe how estimation of the unknown functions in the log-additive temperature model can be formulated as a large linear regression problem. In Ref. 12, we show how a smoothness penalty function can be used to advantage in the spline fit to the additive log-temperature model. We denote the fitted response function by $\hat{T}(\bar{\psi}, \mathbf{u} | f_0, \dots, f_L)$. The algorithm of Ref. 12 is simply: Minimize with respect to the B-spline coefficients of $f_0(\bar{\psi}), \dots, f_L(\bar{\psi})$ of Eq. (4), the weighted least squares problem:

$$\sum_{i,j} \left| \frac{T_i(\bar{\psi}_j^i) - \hat{T}(\bar{\psi}_j^i, \mathbf{u}_i | f_0, \dots, f_L)}{\sigma_{i,j}} \right|^2 + \sum_{\ell=0}^L \lambda_\ell \int_0^1 |f_\ell'''(\bar{\psi})|^2 d\bar{\psi}, \quad (6)$$

where $T_i(\bar{\psi}_j^i)$ is the j th radial measurement of the i th measured temperature profile and $\sigma_{i,j}$ is the associated error. The second term is the smoothness penalty which damps artificial oscillations in the estimated $f_\ell(\bar{\psi})$. The smoothing parameter, λ_ℓ , controls the smoothness of the estimate of $f_\ell(\bar{\psi})$. The Appendix describes how we determine λ_ℓ empirically.

There are two types of systematic error: model error (since the additive model is only an approximation), and smoothing error from the smoothness penalty function. Simplifying the

physical model causes bias error, but can often reduce the variance of the fitted model. This variance reduction occurs because the simplified model usually has fewer free parameters than a more complete model does.

We wish to choose the additive model which minimizes the error in predicting the temperature of a new profile. Unfortunately, the prediction error depends on the unknown “true” temperature function. To select the best model and smoothing parameters, the expected average square error (EASE) is estimated empirically^{14,15}. We use a generalization of the estimate of the EASE given by the Rice criterion. (See the Appendix.) This EASE estimate includes the bias error associated with the incomplete model, i.e. we admit that our additive model is systematically wrong, and we estimate the size of this error.

From this estimate of the EASE, we then select the additive model which minimizes the Rice criterion. Similarly, we choose the smoothing parameters, λ_ℓ , to minimize this empirical estimate of the expected error. The Rice criterion estimates the fit error for *new* data while the older “ χ^2 ” statistic considers the fit quality for the existing data set. The Rice criterion is more selective than the χ^2 statistic in the sense that it prefers simpler, lower order models.

III. JET OHMIC TEMPERATURE PROFILE PARAMETERIZATION

a) Single Profile Analysis

We consider a 43 profile data set from the Joint European Torus¹³. The electron temperature and density profiles are measured by the JET LIDAR Thomson scattering diagnostic. Each profile is measured at approximately 50 radial locations along the plasma mid-plane. Table 1 summarizes the global parameters of the data. The data contains discharges with the edge safety factor, q_{95} as high as 12.

The JET discharges were produced between 1989-90 and 1991-92. During this time, JET operated with carbon tiles, with carbon tile and beryllium evaporation, and with beryllium tiles. Most of the discharges in the database have the plasma boundary formed by the outer wall limiter with beryllium tiles.

The JET LIDAR Thomson scattering diagnostic is described in Ref. 16. A number of the profiles have an artificial increase in the measured temperature near the inner wall. This problem occurs because the laser light from the LIDAR diagnostic is partially reflected near the plasma wall and stimulates radiation emission near the plasma edge. To prevent these spurious data from influencing the profile fit, we delete measurements in the outer ten percent of the plasma which have the temperature increasing near the wall.

Near the inboard wall, there are only rarely usable measurements. To be able to estimate the temperature for $\bar{\psi} < -0.87$, we reflect the temperature for $\bar{\psi} > +0.87$. In our model selection criterion, we use the number of measured data points and not the number of augmented data.

Var	mean	min	max	std dev
\bar{n}	2.29	1.32	3.90	0.75
q95	5.48	2.88	12.6	2.86
I_p	2.80	0.97	5.25	1.13
B_t	2.76	1.30	3.22	0.46
κ	1.44	1.30	1.75	0.122
a	1.16	1.05	1.19	0.040
R	2.92	2.83	3.01	0.047
V_{loop}	-.35	-1.12	.914	0.66
$Z_{eff,1}$	1.88	1.07	3.10	0.55
$Z_{eff,2}$	2.10	1.20	3.35	0.60

Table 1: Database Summary: Average, minimum, maximum and standard deviation of each of the engineering variables.

Neither the fitted functions, $\hat{f}_\ell(\bar{\psi})$, nor the measured data are symmetric with respect to $\bar{\psi}$. The measured profiles are clearly hotter and broader at the inboard side (negative $\bar{\psi}$). Our fitted profiles allow asymmetry and reproduce this trend. Since the LIDAR measurements near the inner wall tend to be less accurate than those on the outboard side, we prefer to fit the data with an asymmetric profile. If a symmetric fit is desired, we recommend using our fit restricted to $\bar{\psi} \geq 0$. It is unclear if this asymmetry is due to a systematic error in the flux map or in the LIDAR measurements or has an unknown physical cause.

Our fitted profiles allow asymmetry and reproduce this asymmetry. We do test each f_ℓ separately for symmetry. Making $f_1(\bar{\psi}) \dots f_4(\bar{\psi})$ symmetric in $\bar{\psi}$ while keeping $f_0(\bar{\psi})$ asymmetric in our best fit model raises the fit error only slightly from 150 to 152 eV. If we force $f_0(\bar{\psi})$ to be symmetric, the difference in the residual fit error is noticeable.

In our spline fits, we use 20 knots. To reduce the ill-conditioning of the fit near the plasma edge, we decrease the density of knots near the edge. Our profile fits depend only very weakly on the knot spacing due to the smoothness penalty term. In contrast, if no smoothness penalty is used (as in the original algorithm of Refs. 10-11), the fit is strongly influenced by knot selection. Our data fit gives an accurate fit to the data which is nearly independent of the knot positions. The smoothing spline yields accurate representation of the solution with fewer artificial oscillations than the methodology of Ref. 10.

Fitting each profile separately gives a root mean square error (RMSE) of .171 on the logarithmic scale, which corresponds to a relative fit error of 17.1 %. On the linear scale the average fit error is 152 eV which is 10.5 % percent of the typical line average temperature (1.454 KeV). The root mean square error (RMSE) is much larger (187 eV) than the mean absolute error, indicating that a small percentage of the data points are very poorly. For the

individual fits, the Rice criterion selects a relatively small amount of smoothing and the fit tends to follow the small scale oscillations in the data.

b) Model Selection

We begin by considering the profile consistency model of Eq. (2):

$$\ln[T] = f_0(\bar{\psi}) + c_I \ln[I_p] + c_B \ln[B_t] + c_n \ln[\bar{n}] + c_\kappa \ln[\kappa] + c_a \ln[a]. \quad (7)$$

This profile consistency model has a Rice criterion value of 1.26. This is 42% larger than our final nonparametric fit in Eq. (8). Thus, the spatial dependencies $\ln[I_p]$, $\ln[B_t]$ and $\ln[\bar{n}]$ are significant. We now consider nonparametric models which include spatial variation in the control variable dependencies.

To select our log-linear model, we use a selection procedure based on the Rice criterion. Our list of candidate variables is $\ln[I_p]$, $\ln[B_t]$, $\ln[\bar{n}]$, $\ln[q_{95}]$, $\ln[\kappa]$, Z_{eff} , V_{loop} , a , R , ℓ_i , and time. At the ℓ -th stages, we try all possible combinations of the best model at the $(\ell - 1)$ -th stage plus one additional variable. We choose the model which reduces the Rice criterion the most.

Table 2 summarizes the reduction in the Rice criterion at each stage. $\ln[I_p]$ is clearly the most important control variable. This result contrasts with earlier profile consistency studies^{12,17} which found that the edge safety factor, q_{95} , is the most important variable in determining the temperature profile shape. This difference occurs because we are fitting the unnormalized temperature and I_p is much more important than q_{95} in determining the magnitude of the temperature.

At the second stage of the sequential selection procedure, the pair $(\ln[I_p], \ln[B_t])$ minimizes the Rice criterion. In the second column of Table 2, we compare the two control variable model: using $f_I(\bar{\psi})\ln[I_p] + f_B(\bar{\psi})\ln[B_t]$ with the two control variable model and using $f_I(\bar{\psi})\ln[I_p] + f_q(\bar{\psi})\ln[q_{95}]$, and we show that the Rice criterion is lower when I_p and B_t are used. The same result holds when $\ln[\bar{n}]$ is added to both models. At the third stage, $(\ln[I_p], \ln[B_t], \ln[\bar{n}])$ has the lowest value of the Rice criterion.

At the fourth stage, $\ln[q_{95}]$ had the lowest value of the Rice criterion when paired with the “seed variables”, $(\ln[I_p], \ln[B_t], \ln[\bar{n}])$. Initially, we had difficulty accepting this result because the I_p and B_t dependencies of q_{95} were already accounted for in the model and the geometric parameters, κ , a and R vary very little. To test if this result were real, we explicitly removed the I_p and B_t dependencies from q_{95} by defining $\hat{q}_{\text{geo}} \equiv q_{95}I_p / B_t$. We found that adding $\ln[\hat{q}_{\text{geo}}]$ resulted in an even smaller value of the Rice criterion than adding $\ln[q_{95}]$. This occurs because $f_{\hat{q}_{\text{geo}}}(\bar{\psi})$ varies less than $f_{q_{95}}(\bar{\psi})$. As a result, we use fewer effective degrees of freedom to represent $f_{\hat{q}_{\text{geo}}}(\bar{\psi})$.

Sequential Selection Using Rice Criterion

Vars in model	1 Var	2 Var	3 Var	4 Var	5 Var
$\ln[\bar{n}]$	4.64	1.73	1.12	seed	seed
$\ln[q_{95}]$	3.04	1.78	1.36	.885	seed
$\ln[I_p]$	1.93	seed	seed	seed	seed
V_{loop}	4.63	1.92	1.53	1.10	.861
$\ln[B_t]$	3.96	1.58	seed	seed	seed
$\ln[\kappa]$	4.30	1.94	1.56	1.10	.875
ℓ_i	4.21	1.63	1.48	1.05	.875
a	4.60	1.91	1.58	1.11	.865
R	4.47	1.88	1.58	1.06	.869
$Z_{eff,1}$	4.01	1.91	1.55	1.06	.872
$Z_{eff,2}$	4.37	1.79	1.57	1.11	.850
Time	4.58	1.87	1.52	.923	.793

Table 2: Rice criterion as a function of the variables in the model. "Seed variables" are included in each run in that column. We then add the variable that reduces the criterion the most.

In general, adding a fifth variable resulted in little reduction in the Rice criterion and the new function, $f_5(\bar{\psi})$, would be nearly constant with large error bars. As a result, we stopped the log-linear expansion with four control variables. Our final model is

$$\begin{aligned} \ln[T] = & f_0(\bar{\psi}) + f_I(\bar{\psi}) \ln[I_p / I_0] + f_B(\bar{\psi}) \ln[B_t / B_0] \\ & + f_n(\bar{\psi}) \ln[\bar{n} / n_0] + f_q(\bar{\psi}) \ln[\hat{q}_{geo} / \hat{q}] \end{aligned} \quad (8)$$

where $\hat{q}_{geo} \equiv q_{95} I_p / B_t$, $I_0 = 2.552$, $B_0 = 2.710$, $n_0 = 2.171$, and $\bar{q} = 4.150$. Table 3 evaluates $f_0(\bar{\psi}) \dots f_4(\bar{\psi})$ at equi-spaced intervals. At the cost of increasing the fit error by 6 % (from 150 to 159 eV), we can replace the model of Eq. (8) with the simpler model

$$\begin{aligned} \ln[T] = & f_0(\bar{\psi}) + f_q(\bar{\psi}) \ln[q_{95} / 4.537] + c_I \ln[I_p / I_0] \\ & + c_B \ln[B_t / B_0] + c_n \ln[\bar{n} / n_0] \end{aligned} \quad (9)$$

where c_I , c_B and c_n are independent of $\bar{\psi}$. The best fit values are $c_I=0.69$, $c_B=0.49$, $c_n=-.37$ and $f_q(\bar{\psi}) \approx -0.2$ for $|\bar{\psi}| < 0.66$ and decreasing to -0.37 at the edge. This simpler model may be more robust than the best fit model of Eq. (8). Table 4 evaluates the coefficients in this model.

c) Fit Results

Figure 1 plots $(f_0(\bar{\psi}))$ and the $f_\ell(\bar{\psi})$. $\exp(f_0(\bar{\psi}))$ is the predicted temperature at $I_p=I_0$, $B_T=B_0$ etc. $f_I(\bar{\psi})$ shows that the temperature broadens and becomes somewhat hollow with increasing current, while $f_B(\bar{\psi})$ shows the same effect with decreasing toroidal magnetic fields. If $f_B(\bar{\psi}) = c - f_I(\bar{\psi})$, then the shape of the profile would depend only on the ratio, B_t/I_p . Thus the shape depends primarily, but not exclusively on B_t/I_p . $f_B(\bar{\psi})$ is more peaked than $f_I(\bar{\psi})$ is hollow, which shows that a relative change in B_t changes the shape more than the corresponding change in I_p . The sequential selection procedure selected I_p over B_t in the first step because I_p varied more than B_t . Thus using I_p reduced the fit error more. $f_{\hat{q}_{geo}}(\bar{\psi})$ is less peaked than either $f_B(\bar{\psi})$ or $f_I(\bar{\psi})$, and therefore changing q_{95} by changing the geometry (a , R and κ) only weakly changes the profile shape.

$f_n(\bar{\psi})$ and $f_{\hat{q}_{geo}}(\bar{\psi})$ are roughly constant, which means that \bar{n} and \hat{q}_{geo} have little effect on the shape of the temperature profile. When $f_n(\bar{\psi})$ and $f_{\hat{q}_{geo}}(\bar{\psi})$ are replaced by constants, the mean absolute residual fit error increases from 150 eV to 156 eV. These constants are also plotted in Figure 1. Figure 2a plots the fitted temperature versus $\bar{\psi}$ and I_p at fixed values of the other parameters. Figure 2b shows how the fitted temperature varies with B_t .

Our results differ from earlier ‘‘profile consistency’’ results because we fit both the temperature shape and magnitude simultaneously; i.e. we do not normalize the data. The sequential selection procedure shows that the total plasma current is more important than the edge safety factor in determining the JET Ohmic temperature profile. From the shape of the $f_\ell(\bar{\psi})$, we see that the polynomial models of the radial dependence poorly approximate the actual shape

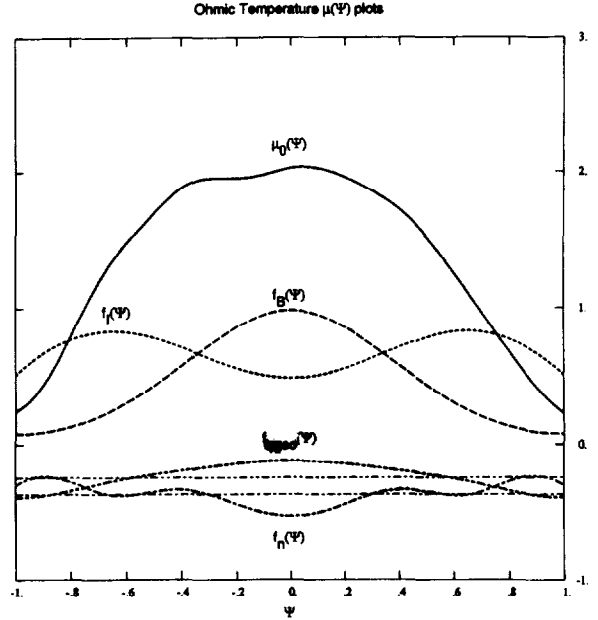


Fig.1: Mean profile, $\mu_0(\bar{\psi})$, and control variable profiles of the log-linear model: $f_I(\bar{\psi})$ through and $f_A(\bar{\psi})$ in Eq. (8). $f_I(\bar{\psi})$ and $f_B(\bar{\psi})$ show that the temperature broadens and becomes somewhat hollow with increasing current or decreasing toroidal magnetic fields. Both $f_n(\bar{\psi})$ and $f_{\hat{q}_{geo}}(\bar{\psi})$ are constant within the error bars of our estimate indicating that \bar{n} and \hat{q}_{geo} have no significant influence on the shape of the temperature profile given the shape modifications of I_p and B_t .

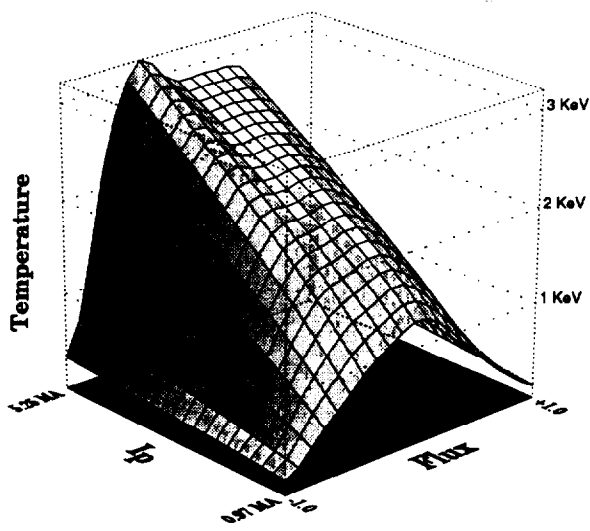


Fig.2a: Surface plot of the predicted Ohmic temperature profile versus I_p evaluated at $B_t=2.710$ T, $\bar{n} = 2.1712 \times 10^{19}$, and $\hat{q}_{geo} = 4.273$.

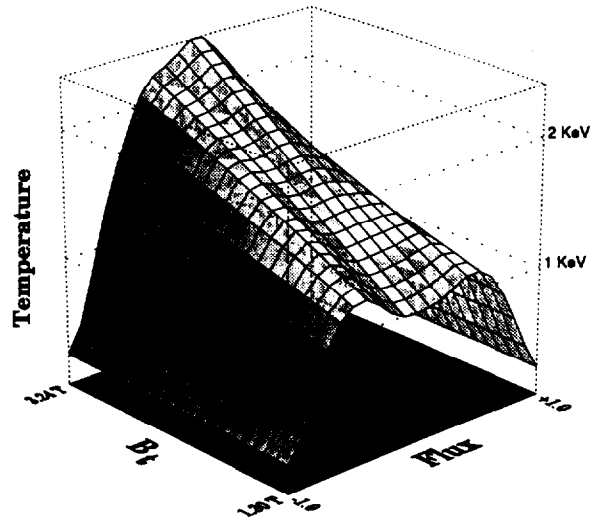


Fig.2b: Surface plot of the predicted Ohmic temperature profile versus B_t evaluated at $I_p=2.552$, $B_0=2.710$, $\bar{n}=2.1712 \times 10^{19}$, and $\hat{q}_{geo} = 4.273$.

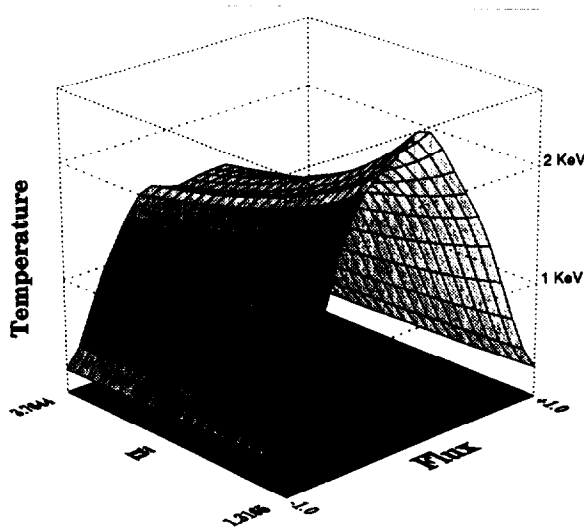


Fig.2c: Response surface of the Ohmic temperature profile versus \bar{n}

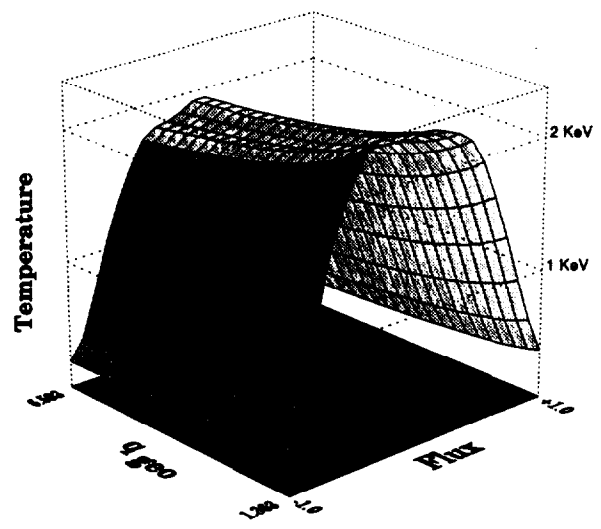


Fig.2d: Response surface of the Ohmic temperature profile versus $q_{95} I_p / B_t$

Figure 3 plots a number of the fitted profiles to illustrate the goodness of fit. Our fitted curve is generally inside the experimental error bars. The combined fit to Eq. (8) gives a mean absolute error (RMSE) of 152 eV. Since the mean line averaged temperature is 1.454 KeV, this

is a 10.5 % typical error. On the logarithmic scale, the RMSE is .171, which corresponds to a relative fit error of 17.1 %.

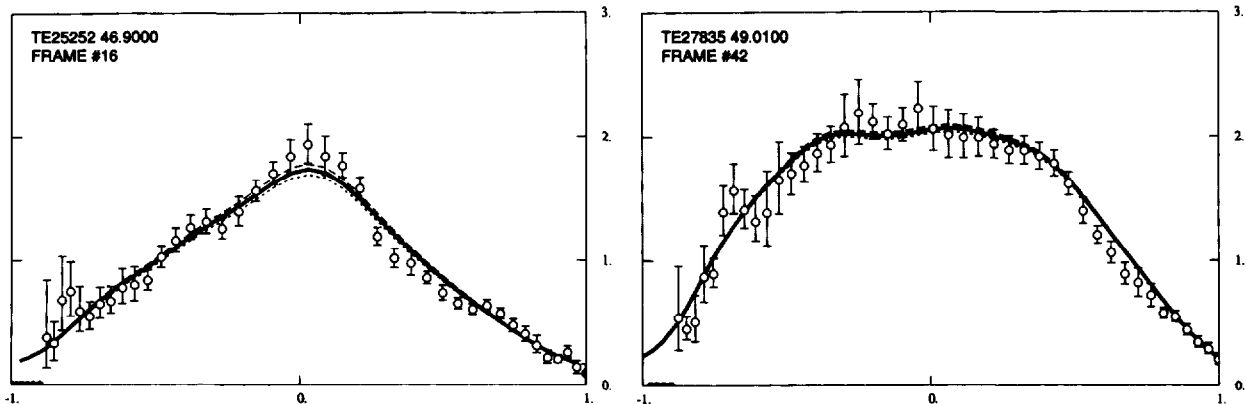


Fig.3a: Log-linear fit to the temperature profile of discharge 25252, time=46.9s, using the the model of Eq. (8).

Fig.3b: Log-linear fit of discharge 27835, time=49.0s.

The mean square error is relevant when the errors have a Gaussian distribution. In our fit, a small percentage of the data has much larger residual errors than is typical. Averaging the square error instead of averaging the absolute error inflates the influence of poorly fitting points. We believe that the mean absolute residual is a more relevant description of the quality of fit.

The Rice criterion value of 0.88 usually means that the expected square error in predicting new data is 0.88 times larger than the experimental variance. For independent errors, the Rice value should be greater than one. Due to the oversampling of the LIDAR diagnostic, the measurement errors are autocorrelated, and we are able to fit the data with smaller residual fit error. In the Appendix, we derive a correction for the autocorrelation. Nevertheless, we ascribe the smallness of our Rice value to the spatial autocorrelation and possibly to uncertainties in the experimental error bars. The Rice value of 0.88 is surprisingly small, given the simplicity of our model and the diverse set of plasma conditions in the database. Thus, we consider this small enhancement to be a major success.

A surprising result of our analysis is that ℓ_i is not particularly useful in estimating the temperature. ℓ_i is the measured value of the second moment of the poloidal magnetic field. If we assume that the current distribution is given by Spitzer resistivity (with a constant spatially, uniform Z_{eff} profile) then ℓ_i can be related to a spatially weighted moment of the temperature distribution. Thus we would expect that larger values of ℓ_i , corresponding to peaked current profiles, would correlate with peaked temperature profiles. Our empirical observation of only a weak dependence of the temperature shape on ℓ_i shows that the current and temperature profile

shapes are partially decoupled. This could be due to variation in Z_{eff} or due to the empirical resistivity differing from the Spitzer value.

At the fifth stage, “time”, as measured from the beginning of the discharge, is the next most important variable. We would like to restrict our analysis to time points in the flat-top. However, there are time points in the early phase of the current ramp down. Adding a time variable to our regression analysis corresponds to the $\ln[T](\bar{\psi}, t) \sim f_0(\bar{\psi}) + f_t(\bar{\psi})(t - \bar{t})$. Our estimate of $f_t(\bar{\psi})$ shows that earlier times tend to be more peaked and later times are flatter. More specifically, $f_{\text{time}}(\bar{\psi})$ strongly resembles $f_1(\bar{\psi})$, which indicates that the profile shape is influenced by the time history of the plasma current. Although our log-additive model is designed for the steady state part of the discharge, most of these discharges have a similar time history. This explains why a nonphysical variable like “time” could reduce the fit error. One of the principal disadvantages of our log-linear temperature models is that the profile shape does not adjust in a physical manner when different time evolution scenarios are used. Thus, $f_t(\bar{\psi})$ is an artefact of the standard time history scenario in JET, and we reject using time as a control variable.

The electron temperature tends to be hotter with carbon tiles than with beryllium tiles on the limiter. Our empirical fit given in Eq.(8) fits both classes of discharges, but the fit parameters were basically determined by the 37 beryllium discharges. Given a larger data set, we could quantify the systematic differences between carbon and beryllium. There are three divertor discharges in our database. The other 40 profiles are limited by the outer wall. Thus the free functions in our empirical fit of Eq.(8) are determined primarily by the limiter discharges. We caution that our results are based on a limited database of 43 JET profiles and that other subsets of the JET data could show different dependencies.

IV. SEMIPARAMETRIC MODELS OF THE DIFFUSIVITY

Much effort has been devoted to determining the anomalous heat diffusivity as a function of the local variables. Hundreds of anomalous transport models have been proposed, none of which is widely accepted. In contrast, there is a consensus of what the “stereotypical” heat diffusivity looks like: the heat diffusivity is usually flat in the inner half of the plasma radius and then increases parabolically in the outer half of the plasma. Furthermore, the radial variation of the heat diffusivity profile appears to depend only weakly on the plasma parameters. There are exceptions to this general assertion, but we believe that broad characterization of the heat diffusivity profile has been supported by many empirical studies.

In constructing empirical models of the anomalous heat diffusivity, our basic hypothesis is that the single most important variable for plasma transport analysis is the normalized flux radius. In other words, normalized flux radius is a more important control variable than more physical variables such as the poloidal gyro-radius. This assertion is difficult to prove or

disprove, so we content ourselves with describing a family of models which are based on this hypothesis. We also wish to parameterize the observed heat diffusivity as a function of the engineering variables in order to influence design studies.

Thus, we propose a second class of models which is similar to the temperature models except that we model the log-diffusivity. A “diffusivity consistent” model is:

$$\ln[\chi(\bar{\psi}, \mathbf{u})] = g_0(\bar{\psi}) + H(\mathbf{u}). \quad (10)$$

Equation (10) implies that as a result the shape of χ depends solely on radius and that the shape of diffusivity is independent of both the local and global plasma parameters including the temperature gradient. Thus, the model is Bohm-like for radial variation. We believe that models similar to Eq. (10) approximate the experimental data fairly well in the sense that the diffusivity tends to be flat out to $\bar{\psi} = .6$, and then increases parabolically. We have not yet modelled plasma discharges by parameterizing the diffusivity, but similar models have been used in plasma modelling¹⁹.

In general, the radial variation of $\ln[\chi]$ can be represented as a slowly varying function of flux radius and the engineering variables. Therefore, we generalize Eq. (10) to all log-additive models of diffusivity

$$\ln[\chi(\bar{\psi}, \mathbf{u})] = g_0(\bar{\psi}) + \sum_{\ell=1}^L g_{\ell}(\bar{\psi})h_{\ell}(\mathbf{u}) \quad (11)$$

The inward density pinch can be included with a similar linear model. As in Eq. (4), we assume that the $h_{\ell}(\mathbf{u})$ are known and are typically $\ln[I_p]$, $\ln[\bar{n}]$ and $\ln[B_t]$. The $g_{\ell}(\bar{\psi})$ are usually given by other smoothing splines or low order polynomials.

The physics implications of the two classes of models - additive log-temperature models and additive log-diffusivity models - are different. The additive log-diffusivity models adjust the temperature profile shape as the radial distribution of sinks and sources. In contrast, the additive log-diffusivity model predicts that the temperature profile shape does not depend only on the global parameters and not on the radial distribution of sinks and sources.

Profile resilience can be interpreted as the observation that the truth is somewhere between these two viewpoints. In other words, the temperature profile shape adjusts less than one would expect from a diffusive model¹⁸. In future work, we hope to compare the two classes of models to see whether additive shape models better describe the temperature on the diffusivity.

We can add additional variables to the control variable vector, \mathbf{u} , such as the beam penetration depth normalized to the minor radius, which partially specify the heating profile. In this way, we can have additive log-temperature models adjust to heating profiles and additive log-diffusivity models be more profile resilient.

In the next section, we discuss the underlying difficulties in parameterizing the diffusivity.

V. ESTIMATION OF THE ADDITIVE LOG-DIFFUSIVITY MODEL

To estimate the B-spline coefficients for the additive log-diffusivity model, we minimize Eq.(5) as well. For the additive log-temperature model, the predicted values, $\hat{T}(\bar{\psi}_j^i, \mathbf{u}_i | g_\ell)$, are a linear function of the spline coefficients and the resulting functional is quadratic. In contrast, in the additive log-diffusivity model, $\hat{T}(\bar{\psi}, \mathbf{u} | g_\ell)$, is a nonlinear function of the unknown spline coefficient and each evaluation of $\hat{T}(\bar{\psi}, \mathbf{u} | g_\ell)$, in the minimization of Eq.(6) requires the solution of the transport equation using a code such as SNAP²⁰ or TRANSP. The sinks and sources may be calculated for each discharge separately prior to beginning the least squares fit for the additive log-diffusivity model. Thus the evaluation of $\hat{T}(\bar{\psi}, \mathbf{u} | g_\ell)$, requires only the inversion of a heat transport equation with *fixed sinks and sources* at each step of the minimization of Eq.(6).

We are unaware of any local χ /heat flux regression study that has attempted to use $\bar{\psi}$ and the global engineering variables with a model similar to Eq. (11). Instead, researchers have tried to determine the dependencies of the diffusivity on heat flux by regressing the point estimates of χ or $\chi \nabla T$ versus local quantities. Previous local χ /heat flux regressions have ignored $\bar{\psi}$ and have concentrated on local quantities such as the poloidal gyro-radius. This approach has several disadvantages relative to the additive log-diffusivity approach with a global minimization. First, we believe that the most useful variable in fitting the diffusivity shapes is the normalized plasma radius, $\bar{\psi}$. Second, we believe that usually $\ln[\chi]$ has a simple and smoothly varying radial dependence. Third, the errors in estimates of the plasma gradients are often comparable to the errors in χ and larger than the errors in the heat flux. When the dependent variables have errors, linear regression is an inconsistent estimator of the parameters. Even worse, the errors of the dependent and independent variables are strongly correlated. Our personal experience is that when χ is regressed against ∇T , the most likely result is that $\chi \sim \frac{I_p \nabla_{loop}}{n \nabla T}$. This result is strikingly similar to the definition of χ . Note the similarity of this expression and the Coppi-Gruber-Mazzucato formula²¹.

In the last paragraph, we described the problems in regressing the local heat flux/ χ versus the local plasma parameters instead of radius and the global plasma parameters. We now describe a second set of disadvantages which persist even when the independent variables are $\bar{\psi}$ and \mathbf{u} . First, by fitting with Eq. (10), we force $\chi(\bar{\psi}, \mathbf{u})$ to be a smoothly varying function. If the estimated χ is regressed at each radial point separately, $\hat{\chi}$ will usually have spurious spatial oscillations. Another severe disadvantage of regressing the local heat flux/ χ directly is that the heat flux is measured and not inferred. Changing $\hat{\chi}(\bar{\psi}, \mathbf{u})$ at one spatial location will modify the predicted temperature, \hat{T} , at all radial locations. Thus to attain a self-consistent estimation of χ , we are forced to fit all radial locations simultaneously. Finally, estimates of the variance of the point estimate of $\hat{\chi}_i(\bar{\psi}_j^i)$ are difficult to obtain, and this makes a pointwise weighted least squares analysis usually infeasible.

For all these reasons, we prefer the global minimization approach of Eq.(5) with the simple additive model of Eq.(11). Nevertheless, the computational and programming effort to fit the log-additive χ model of Eq.(11) is considerable and we have not yet applied it to JET data.

VI. SUMMARY

Profile resilience and diffusivity profile resilience strongly suggest that the appropriate empirical models for local profile dependencies are the additive log-temperature model and additive log-diffusivity model. We therefore distinguish four classes of empirical transport models:

- 1) *Global confinement models*: τ_E (engineering variables) as typified by Eq. (1);
- 2) *Semiparametric profile models*: $T(\bar{\psi}, \text{engineering variables})$ as typified by Eq. (8);
- 3) *Semiparametric diffusivity models*: $\chi(\bar{\psi}, \text{engineering variables})$ as typified by Eq. (11);
- 4) *First principles transport models*: $T(\text{physics variables})$, possibly given by a theoretical expression.

The huge multiplicity of transport theories and the relative lack of success in applying theory based models motivates us to consider the semiparametric models of classes 2) and 3). We hope to use these same models to predict the profile peakedness factor for Deuterium-Tritium (D-T) discharges in JET and the Tokamak Fusion Test Reactor²² (TFTR) and for extrapolating performance to International Thermonuclear Experimental Reactor²³ (ITER).

We have accurately parameterized the JET Ohmic temperature profiles using a log-linear temperature model (Class 2). We have not yet fitted log-additive diffusivity models to the JET data (Class 3). In both cases, the smoothing spline coefficients are best determined by minimizing the residual fit error over all measured profiles simultaneously.

Our parameterized temperature model Eq.(8) fits our JET data set with a mean absolute error of 47 eV which is 3.2 % percent of the typical line average temperature. We recommend using this parameterization for transport analyses and most other applications.

Is the database big enough to make these conclusions? Some of our conclusion will depend on the choice of data. If the database were to contain only discharges from a specialized scan on one or two consecutive days, we would probably be able to fit the data better, including detailed profile features. Because our database is taken over a diverse set of discharge conditions, we average over the small scale features which depend on the particular discharge conditions. As a result, the fit is worse, but the results are more robust because they reflect many different types of discharges.

The Rice criterion measures the expected error in predicting new data, normalized to the variance of the measurements for independent errors. Our Rice value of 0.88 means that our predictions are theoretically more accurate than the experimental measurements. This small value of the Rice criterion is probably due to the spatial autocorrelation from oversampling and possibly to uncertainties in the experimental error bars. We consider this small value of the predictive error to be a major success given the diverse set of plasma conditions in the database and the simplicity of our model.

Fitted Functions for Eq. (8)

$\bar{\psi}$	$f_0(\bar{\psi})$	$f_I(\bar{\psi})$	$f_B(\bar{\psi})$	$f_n(\bar{\psi})$	$f_{\hat{q}_{geo}}(\bar{\psi})$
-1.0	0.2376	0.5057	0.0776	-0.3013	-0.3879
-0.9	0.4267	0.6679	0.0900	-0.2332	-0.3755
-0.8	0.7972	0.7728	0.1320	-0.2710	-0.3370
-0.7	1.1884	0.8231	0.2037	-0.3479	-0.2902
-0.6	1.4813	0.8236	0.3046	-0.3746	-0.2484
-0.5	1.7071	0.7827	0.4317	-0.3449	-0.2161
-0.4	1.8869	0.7131	0.5771	-0.3294	-0.1891
-0.3	1.9517	0.6316	0.7261	-0.3658	-0.1631
-0.2	1.9528	0.5561	0.8577	-0.4384	-0.1397
-0.1	1.9785	0.5029	0.9491	-0.5021	-0.1236
0.0	2.0271	0.4838	0.9820	-0.5261	-0.1179
0.1	2.0257	0.5029	0.9491	-0.5021	-0.1236
0.2	1.9588	0.5561	0.8577	-0.4384	-0.1397
0.3	1.8611	0.6316	0.7261	-0.3658	-0.1631
0.4	1.7285	0.7131	0.5771	-0.3294	-0.1891
0.5	1.5236	0.7827	0.4317	-0.3449	-0.2161
0.6	1.2578	0.8236	0.3046	-0.3746	-0.2484
0.7	0.9711	0.8231	0.2037	-0.3479	-0.2902
0.8	0.6821	0.7728	0.1320	-0.2710	-0.3370
0.9	0.4173	0.6679	0.0900	-0.2332	-0.3755
1.0	0.2243	0.5057	0.0776	-0.3013	-0.3879

Table 3: Evaluation of the radial spline functions in Eq. (8) for the JET data.

Our sequential selection procedure shows that the current is the most important control variable in determining the temperature profile. Previous studies, which have considered the

normalized temperature, have found that the edge safety factor is the most important control variable. Increasing the plasma current results in broader, often hollow profiles. Fig.1 shows that the current and the magnetic field modify the profile shape in different ways. Thus, the temperature shape does not depend exclusively on q_{95} .

The profile parameterization in Eq. (8) is only for time-independent profiles. The log-linear diffusivity model may be more relevant to modelling time evolution. Since our database consist almost exclusively of limiter discharges, we are unable to determine if the temperature profile is modified when a divertor is used.

Fitted Functions for Eq. (9)

$\bar{\psi}$	$f_0(\bar{\psi})$	$f_q(\bar{\psi})$	$f_I(\bar{\psi})$	$f_B(\bar{\psi})$	$f_n(\bar{\psi})$
-1.0	0.2326	-0.3729	0.6868	0.4900	-0.3652
-0.9	0.4279	-0.3364	0.6868	0.4900	-0.3652
-0.8	0.8015	-0.2822	0.6868	0.4900	-0.3652
-0.7	1.1914	-0.2298	0.6868	0.4900	-0.3652
-0.6	1.4827	-0.1951	0.6868	0.4900	-0.3652
-0.5	1.7092	-0.1836	0.6868	0.4900	-0.3652
-0.4	1.8890	-0.1885	0.6868	0.4900	-0.3652
-0.3	1.9514	-0.1995	0.6868	0.4900	-0.3652
-0.2	1.9505	-0.2105	0.6868	0.4900	-0.3652
-0.1	1.9750	-0.2186	0.6868	0.4900	-0.3652
0.0	2.0194	-0.2215	0.6868	0.4900	-0.3652
0.1	2.0154	-0.2186	0.6868	0.4900	-0.3652
0.2	1.9511	-0.2105	0.6868	0.4900	-0.3652
0.3	1.8579	-0.1995	0.6868	0.4900	-0.3652
0.4	1.7298	-0.1885	0.6868	0.4900	-0.3652
0.5	1.5275	-0.1836	0.6868	0.4900	-0.3652
0.6	1.2613	-0.1951	0.6868	0.4900	-0.3652
0.7	0.9729	-0.2298	0.6868	0.4900	-0.3652
0.8	0.6841	-0.2822	0.6868	0.4900	-0.3652
0.9	0.4194	-0.3364	0.6868	0.4900	-0.3652
1.0	0.2208	-0.3729	0.6868	0.4900	-0.3652

Table 4: Evaluation of the radial spline functions in Eq. (9) for the JET data.

**APPENDIX:
RISK ESTIMATION AND MODEL SELECTION**

The estimation of risk/expected error is critical to our analysis because we use this estimate to select which terms to include in our analysis. We represent the “true” log-temperature values by the vector μ and the measured log-temperature by $y \equiv \mu + \varepsilon$. We present the generalized cross-validation (GCV) estimate of the expected average square error (EASE) as well as the Rice criterion correction. We consider the linear regression model:

$$y \equiv \mu + \varepsilon, \quad \mu = X\alpha, \quad (A1)$$

where y is the measurement vector, X is the data matrix, α is the parameter vector and ε is a vector of random errors with covariance matrix Σ . We define D to be a diagonal matrix which contains the inverses of the variances of the measurements: $D_{i,j} = \sum_{i,i}^{-1} \delta_{i,j}$. Presently, we do not include the off-diagonal terms in Σ in the minimization, but do compensate for this in our model selection criterion.

By $\mu = X\alpha$, we mean that we model μ by $X\alpha$, but that we admit that $\mu = X\alpha$ is not exact and that this model has a systematic error. We estimate α using the penalized least squares estimate:

$$\hat{\alpha}_\lambda = \arg \min_{\alpha} \left\{ (y - X\alpha)^\dagger D (y - X\alpha) + \lambda \alpha^\dagger S \alpha \right\}, \quad (A2)$$

where S is the penalty matrix. Equation (A2) is an abstract matrix formulation of Eq.(4). For brevity, we denote $z^\dagger D z$ by $\|z\|_D^2$, where z is an arbitrary n -vector.

The subscript λ on $\hat{\alpha}_\lambda$ denotes the dependence on the smoothing parameter. Equation (A2) can be rewritten as

$$\hat{\alpha}_\lambda = [X^\dagger D X + \lambda S]^{-1} X^\dagger D y = G_\lambda X^\dagger D y, \quad (A3)$$

where $G_\lambda \equiv [X^\dagger D X + \lambda S]^{-1}$. The covariance $\hat{\alpha}_\lambda$ is

$$\text{Cov}[\hat{\alpha}_\lambda \hat{\alpha}_\lambda^\dagger] = G_\lambda K G_\lambda \quad (A4)$$

where $K \equiv X^\dagger D \Sigma D X$. In addition to the variance, Eq. (A3) has a bias/systematic error: $E[\mu - X\hat{\alpha}_\lambda]$. The dominant source of bias error in our analysis is due to model error in the additive model. The expected average square error (EASE) in the fit is

$$EASE = E\left[\|\mu - X\hat{\alpha}_\lambda\|_D^2\right] = \text{Bias}^2 + \text{Variance} = \|E[\mu - X\hat{\alpha}_\lambda]\|_D^2 + \text{trace}[C G_\lambda K G_\lambda], \quad (A5)$$

where $\mathbf{C} \equiv \mathbf{X}^\dagger \mathbf{D} \mathbf{X}$ and $\text{trace}[\mathbf{C} \mathbf{G}_\lambda \mathbf{K} \mathbf{G}_\lambda] = \text{trace}[\mathbf{D} \mathbf{X} \text{Cov}[\hat{\alpha}_\lambda \hat{\alpha}_\lambda^\dagger] \mathbf{X}^\dagger]$.

We wish to minimize the EASE. However, it is unknown and needs to be estimated. We denote the average square residual of the empirical fit by $\hat{\sigma}_\lambda^2(\lambda) \equiv \|\mathbf{y} - \mathbf{X} \hat{\alpha}_\lambda\|_{\mathbf{D}}^2 / N$. The expectation of the square residual error is

$$\mathbb{E}[\|\mathbf{y} - \mathbf{X} \hat{\alpha}_\lambda\|_{\mathbf{D}}^2] = \text{Bias}^2 + \{\text{trace}[\mathbf{\Sigma} \mathbf{D}] - 2 \text{trace}[\mathbf{K} \mathbf{G}_\lambda] + \text{trace}[\mathbf{C} \mathbf{G}_\lambda \mathbf{K} \mathbf{G}_\lambda]\}. \quad (\text{A6})$$

Note $\text{trace}[\mathbf{\Sigma} \mathbf{D}] = N$. Equation (A5) computes the error relative to the true, unmeasured values while Eq. (A6) uses the measured residuals. As a result, Equation (A5) can be easily estimated from the data by computing the MSE of the fit using the measured temperature. In contrast, Eq. (A5) involves the unknown, “true” temperature. The Craven-Wahba estimate of the EASE uses Eq. (A6) to estimate Eq. (A5):

$$\mathbb{E}[\|\mu - \mathbf{X} \hat{\alpha}_\lambda\|_{\mathbf{D}}^2] = \|\mathbf{y} - \mathbf{X} \hat{\alpha}_\lambda\|_{\mathbf{D}}^2 - \{N - 2 \text{trace}[\mathbf{K} \mathbf{G}_\lambda]\}. \quad (\text{A7})$$

Equation (A7) is particularly valuable because it includes the systematic error in directions which are orthogonal to the column space of \mathbf{X} ; i.e. the bias from not including all possible terms in the additive model. At $\lambda=0$, $\text{trace}[\mathbf{C} \mathbf{G}_\lambda]$ equals the number of fit parameters and decreases monotonically with λ . Using \mathbf{K} in place of \mathbf{C} compensate for the autocorrelation of the measurements.

The minimum of Eq. (A7) with respect to λ satisfies:

$$\partial_\lambda \hat{\sigma}_\lambda^2 + \frac{2}{N} \partial_\lambda \text{trace}[\mathbf{K} \mathbf{G}_\lambda] = 0 \quad (\text{A8})$$

Equation (A8) is useful when σ^2 is known. We now consider the case where the covariance of the measurements is known up to an arbitrary constant: $\text{Cov}[\mathbf{y} \mathbf{y}^\dagger] = \sigma^2 \mathbf{\Sigma}$, where σ^2 is unknown and $\mathbf{\Sigma}$ is known. We continue to define $\mathbf{D}_{i,j} = \sum_i \delta_{i,i}^{-1} \delta_{i,j}$ and keep the same definitions for \mathbf{G}_λ , \mathbf{C} and \mathbf{K} . When σ^2 is unknown, we can estimate it using the Craven-Wahba estimate of σ^2 :

$$\hat{\sigma}_{CW}^2 = \frac{\|\mathbf{y} - \mathbf{X} \hat{\alpha}_\lambda\|_{\mathbf{D}}^2}{(N - \text{trace}[\mathbf{K} \mathbf{G}_\lambda])} \quad (\text{A9})$$

In penalized regression, “ $N - \text{trace}[\mathbf{K} \mathbf{G}_\lambda]$ ” is referred to as the effective number of degrees of freedom.

The empirical estimate of Eq. (A8) using the estimate $\hat{\sigma}_{CW}^2$ is

$$\partial_{\lambda} \hat{\sigma}_{\lambda}^2 + \frac{2\hat{\sigma}_{\lambda}^2}{N - \text{trace}[\mathbf{K}\mathbf{G}_{\lambda}]} \partial_{\lambda} \text{trace}[\mathbf{K}\mathbf{G}_{\lambda}] = 0, \quad (\text{A10})$$

which implies

$$\partial_{\lambda} \left\{ \frac{\hat{\sigma}_{\lambda}^2}{\left(1 - \text{trace}[\mathbf{K}\mathbf{G}_{\lambda}]/N\right)^2} \right\} = 0 \quad (\text{A11})$$

Thus, we define the generalized cross-validation statistic as

$$GCV \equiv N \frac{\|\mathbf{y} - \mathbf{X}\hat{\alpha}_{\lambda}\|_{\mathbf{D}}^2}{\left(N - \text{trace}[\mathbf{K}\mathbf{G}_{\lambda}]\right)^2}. \quad (\text{A12})$$

Minimizing the GCV criterion of Eq. (A12) sometimes undersmooths and tends to pick models with too many free parameters¹⁵. We refer the reader to Ref.15 for an empirical comparison of the Rice criterion and the GCV criterion. Therefore, we replace Eq. (A10) with a modified loss estimator based on the Rice criterion:

$$C_R \equiv \frac{\|\mathbf{y} - \mathbf{X}\hat{\alpha}_{\lambda}\|_{\mathbf{D}}^2}{\left(N - 2\text{trace}[\mathbf{K}\mathbf{G}_{\lambda}]\right)} \quad (\text{A13})$$

C_R has also been normalized to the standard error per data point. C_R differs from Eq. (A12) by terms of $O(\text{trace}[\mathbf{K}\mathbf{G}_{\lambda}]/N)$. In our analysis, we minimize Eq. (A13) with respect to both the choice of control variables in the additive model and the smoothing parameters in a given model. In Ref. 24, a somewhat different autocorrelation correction is derived.

An older statistic is $\chi^2 \equiv \frac{\|\mathbf{y} - \mathbf{X}\hat{\alpha}_{\lambda}\|_{\mathbf{D}}^2}{N - 2\text{trace}[\mathbf{K}\mathbf{G}_{\lambda}]}$, which corresponds to the mean square error per degree of freedom. The χ^2 statistic is useful in optimizing the fit to existing data while the Rice criterion and generalized crossvalidation are useful in minimizing the predictive error for new data. The factor of two in the denominator of C_R results in smoother models and fewer variables in the model.

ACKNOWLEDGEMENTS

G. Cordey's support and encouragement are gratefully acknowledged. We thank C. Gowers, P. Nielsen, K. Thomsen, and D. Muir. KI's work was supported by the U.S. Department of Energy Grants No.DE-FG02-92ER54157. SR's work was supported by the U.S. Department of Energy Grants DE-FG02-86ER-53223 and 91ER54131.

REFERENCES

- [1] K. S. Riedel, Kaye S. M., *Nuclear Fusion* **30** 731 (1990).
- [2] K. S. Riedel, *Nuclear Fusion* **30** 755 (1990).
- [3] Riedel, K. S., *Comments in Plasma Physics and Controlled Fusion* **12** (1989) 279.
- [4] Yushmanov, P., Takizuka, T., Riedel, K. S., et al., *Nuclear Fusion* **30** (1990) 1999.
- [5] K. S. Riedel, *Nuclear Fusion*, **31** 927 (1991).
- [6] Christiansen, J. P., Cordey, J. G., Kardaun, O. J. W. F., and Thomsen, K., *Nuclear Fusion* **31** 2117 (1991).
- [7] Christiansen, J. P., Cordey, J. G., Thomsen, K., et al., *Nuclear Fusion* **32** (1992) 291.
- [8] Redi, M. H., Tang, W. M., Efthimion, P. C., et al., *Nuclear Fusion* **27** 2001 (1987).
- [9] Tang, W. M., *Nuclear Fusion* **26** 1605 (1986).
- [10] McCarthy, P. J., Riedel, K. S., Kardaun, O. J. W. F., Murmann, H., Lackner, K., *Nuclear Fusion* **13** (1991) 1595, also "Scalings and Plasma Profile Parameterisation of ASDEX High Density Ohmic Discharges", Max-Planck-Institut für Plasmaphysik Report No. 5/34.}
- [11] Kardaun, O. J. W. F., Riedel, K. S., McCarthy, P. J., et al. Max-Planck-Institut für Plasma Physik Report No. 5/35, (1990).
- [12] Riedel, K. S. and Imre, K., *Comm. in Statistics* **22** 1795 (1993).
- [13] Rebut, P.H., Bickerton, R.J., Keen, B.E., *Nuclear Fusion* **25** 1011, (1985).
- [14] Craven, P. and Wahba, G., *Numer. Math.* **31** (1979) 377.
- [15] Hardle, W., Hall, P., and Marron, S., *J. Amer. Statist. Assoc.* **83** (1988) 86.
- [16] Salzmann, H., Bundgaard, J., Gadd, A., Gowers, C. et al., *Rev. Sci. Instrum.* **59** (1988) 1451.
- [17] Arunsalam, V. D., Bretz, N. L., Efthimion, P. C., et al., *Nuclear Fusion* **30** (1990) 2111.
- [18] Phillips, C. K., Houlberg, W., Hwang, D., et al., *Plasma Phys. Control. Fusion* **35** (1993) 301.
- [19] Christiansen, J. P., Callen, J. D., Cordey, J. G., Thomsen, K., *Nuclear Fusion* **28** (1988) 817.
- [20] Towner, H.H., Goldston, R.J., Hammett, G.W., Murphy, J.A., Phillips, C.K., Scott, S.D., Zarnstorff, M.C., Smithe, D., *Reviews of Sci. Inst.* **63** (1992).
- [21] Gruber, O., *Nuclear Fusion* **22** (1982) 1349.
- [22] Grove, D.J., Meade, D.M., *Nuclear Fusion* **25** 1167, (1985).
- [23] Post, D., et al, *ITER Physics Basis*, IAEA Publishing, Vienna, (1991).
- [24] Altman, N. S., *J. Amer. Stat. Assoc.* **85** (1990) pp. 749.

