Statistical induction of a thermal transport model based on the transport analyses database

[1,2]M. Yokoyama

[1]National Institute for Fusion Science, National Institutes of Natural Sciences, 322-6 Oroshi, Toki 509-5292, Japan

[2]SOKENDAI, 322-6 Oroshi, Toki 509-5292, Japan

Abstract

A new approach for inducing a thermal transport model, for the ion heat diffusivity as an example, for magnetically confined high-temperature plasmas has been further pursued after its initial proposal in Ref. [1]. It has been based on a statistical approach utilizing the accumulated experimental transport analyses database. Two approaches are described in this paper: (1) placing a priority on reproducing the ion heat diffusivity with higher accuracy for better reproduction of ion temperature profiles, and (2) attempting to acquire a physics interpretation through variable selections and the dependence of the ion heat diffusivity on them. Such progress will foster the study of a practical transport model for the real-time control and the provision for guidance to the parameter dependence to be pursued by large-scale cutting-edge simulations.

Keywords: transport model, ion heat diffusivity, statistical approach, log-linear multivariate regression, Akaike's Information criterion (AIC)

1. Introduction

Many transport models for magnetically-confined high-temperature plasmas have been proposed. These models are conventionally based on first-principle considerations, and/or heuristic deduction from plausible and responsible physics mechanisms for causing transport. Recently, machine-learning (ML) or data-driven approach has been implemented by utilizing accumulated database in experiments and simulations such as Refs. [2,3]. However, it should be concluded that none of these models have provided accurate and satisfactory models, because neither single first-principle equation nor physics phenomenon is enough for describing complicated phenomena in actual confined plasmas. Further, ML techniques often are "black-boxes," which have weaknesses regarding physics interpretations.

In this paper, progress on a statistical approach originally proposed in Ref. [1] is described by extending and improving statistical analyses. The database for this study is a transport analyses database, which can be considered as a mixture of experimental (actual) and light-simulation results, in terms of utilizing experimental measurements of density and temperature profiles, and numerical analyses results on heating deposition profiles for NBI (neutral beam injection). Then the ion heat diffusivity is estimated by assuming a simple diffusive picture, utilizing those profiles: the calculated heating deposition, density and ion temperature gradient.

Here brief description is provided for the comparison between neural network (NN, one of the ML approaches such as used in [2,3]) and the currently employed statistical approach (SA). Both approaches prescribe a set of variables as "inputs" to NN and "explanatory variables" to SA. The NN is trained with providing outputs simultaneously and does not prescribe any of direct functional forms between "inputs" and "outputs". The NN is based on basis activation (or transfer) function (such as sigmoid) and weights on "neurons". Learning capability can be controlled and improved, depending on width and depth of multi-layers. On the other hand, the statistical approach in this Letter does prescribe a functional form (log-linear) with available and plausible variables to perform multivariate regression analysis. In this sense, modelling capability is limited compared to NN, but apparent functional forms can be easily obtained, by which physics interpretation can be considered as described in Sec. 4. It should be noted here that the log-linear form is just an example of prescribed functional forms, and there is a plenty of other possibilities for functional forms, which may increase the modelling capability.

The paper is organized as follows: The previous scaling law is revisited in Sec. 2, in terms of the application of Akaike's Information Criterion (AIC) [4] to confirm the relevance of AIC for statistical analyses in the subsequent sections. In Sec. 3, it is explained how the database for this study is formulated with its basic characteristics on parameter distributions. Section 4 deals with statistical analyses in two folds, the first being for placing a priority on reproducing the ion heat diffusivity (values) with higher accuracy. The second is for attempting to acquire physics interpretations (trends) through variable selections. Finally, summary and discussion are provided in Sec. 5.

2. Revisiting a previous scaling law based on Akaike's Information Criterion

AIC is the measure for the relative quality of statistical models for a given dataset. It has been employed to consider a balance between the complexity in the model and the goodness of fit of the model. The definition of the AIC is given as follows:

$$\mathrm{AIC} = -2 \ln L + 2k,$$

where L is the maximum likelihood, and k is the number of free parameters. Practically, the minimum AIC selects an "optimal" model among many candidates. More precisely, AICc (AIC with a correction for small sample sizes) [5] is used in this paper. It should be noted here that Bayesian Information Criterion (BIC) is also defined similarly, but $2k$ in AIC is to be replaced by $k\ln(n)$, where n is the number of sample size [6]. In the following discussion, AICc values are referred, however, BIC gives the same tendency with almost the same values as those of AICc due to small sample size (~3000).

Before implementing the AIC estimate in the current study, the previous scaling law for the energy confinement time is revisited to confirm its relevance. An example of a scaling law for this examination is the so-called ISS04 [7], which was derived through log-linear multivariate regression based on the energy confinement time database formulated by contributions from several helical devices. The ISS04 scaling law is expressed as

$$\tau_E = 0.134 a^{2.28} R^{0.64} P^{-0.61} n_{\mathrm{e,bar}}^{0.54} B^{0.84} (\iota/2\pi)^{0.41}, \qquad (1)$$

by employing engineering parameters. Here, τ_E, a, R, P, $n_{\mathrm{e,bar}}$, B and $\iota/(2\pi)$ are energy confinement time [s], plasma minor and major radii [m], absorbed heating power [MW], line-averaged electron density [10^{19} m^{-3}], magnetic field strength [T], and rotational transform at the 2/3 of the plasma minor radius. In Ref. [7], the goodness of the candidate models is compared by root mean square error (RMSE) of the fit. Thus, the ISS04 scaling is revisited here by means of AIC, by comparing possible models using the above-mentioned 6 engineering variables.

Figure 1 shows the evolution of AICc value as a function of the number of variables adopted in the model. It clearly indicates that the maximum number of variables, six in this case, gives the minimum AICc value. It means that the model with all six variables (expression (1)) is an optimal model among all the candidates which are possible with these 6 variables. Thus, ISS04 scaling law (expression (1)) is statistically supported also by means of AICc. The AIC is recognized as a useful and powerful measure for selecting reasonable modelling, which is utilized in the subsequent sections. It should be noted that "the minimum AIC" condition is practically interpreted as the one with "the lowest AIC" among *a priori* prepared physically relevant explanatory variables. This applies here and the subsequent sections.

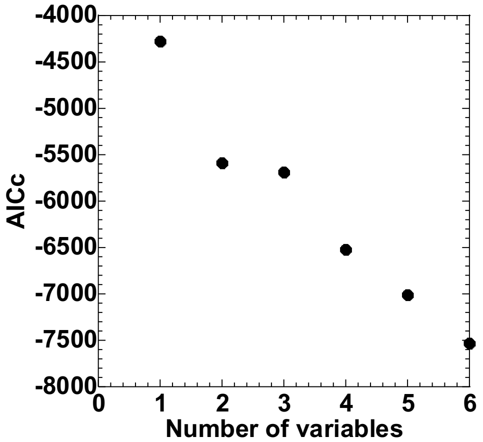

Fig. 1: The evolution of AICc values as a function of the number of variables employed in ISS04 (expression (1)).

3. Description of database for this study

The database considered in this paper is the same as that used in Ref. [1]. For convenience, relevant plots for this database are reproduced in Fig. 2 from Ref. [1]. This is the database (about 3000 data) for the ion thermal diffusivity, χ_i, constructed by considering approximately 200 time slices in 31 discharges in Large Helical Device (LHD) [8]. These 31 discharges are sampled from experiments targeting high ion-temperature (high-T_i) with 2.75 T, and 200 time slices are corresponding to those with ion temperature profile measurement in 31 discharges. The T_i, electron temperature (T_e) and electron density (n_e) at core region range, ~2 to ~7 keV, ~2.5 to ~4 keV, and ~1x to ~1.7x10^{19} m^{-3}, respectively. As recognized in Fig. 2, the range of T_e is smaller than that of T_i, which results in the smaller range of explanatory variables and then poorer statistical properties. Thus, in this letter, let me focus for inducing a model statistically only on ion thermal diffusivity. This database is qualitatively different from that of the energy confinement time, τ_E, in terms of considering all the profiles (ion and electron temperature, and electron density). The database is formulated through the development and extensive applications of the integrated transport analysis suite, TASK3D-a ("a" stands for experimental "A"nalysis) [9] to LHD plasmas. The variables to be employed as explanatory variables for the regression analyses in this paper are all dimensionless. They are summarized in Table 1 with explanations. The plasma beta, one of the usual dimensionless parameters, is not included in Table 1 since it is not used in this paper due to its limited expansion in the database. However, equilibrium changes are appropriately taken into account since transport analyses in TASK3D-a are performed based on experimentally reconstructed (mapped) equilibria by utilizing measured electron temperature profiles (so-called TSMAP system) [10].

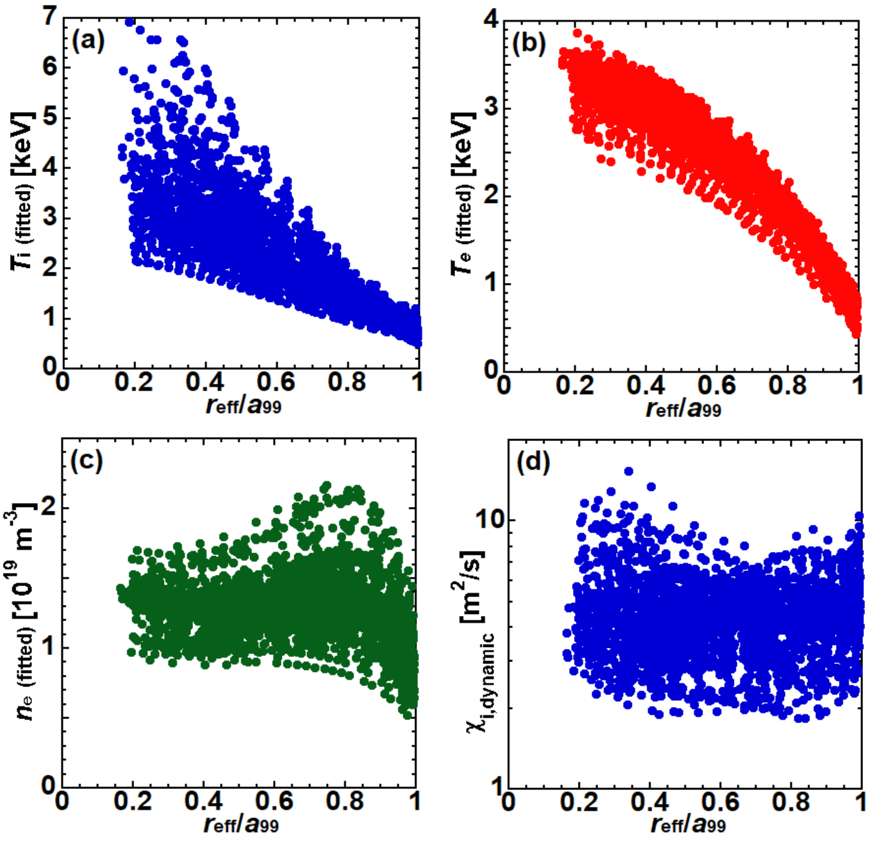

Fig. 2: All the data points of employed database for (a) Ti, (b) Te, (c) ne and then (d) χ_i, which are reproduced from Ref. [1].

Variables	**Explanations**
ν_i*	Ion collision frequency normalized by its banana-plateau boundary
ρ_i*	Ion Larmor radius normalized by plasma minor radius
T_e/T_i	Temperature ratio (electron to ion)
R/L_{Ti}	Normalized inverse scale length of the ion temperature gradient with R being the geometrical center of the outermost flux surface
R/L_{ne}	Normalized inverse scale length of the electron density gradient
$\iota/2\pi$	Rotational transform
ϵ_h	Main helicity of magnetic configuration
ϵ_t	Toroidicity of magnetic configuration
ϵ_{eff}	Effective helicity

Table 1: The list of variables to be employed as explanatory variables for the regression analyses in this paper.

It should be noted again that all of these variables are local values (that is, taking profiles into consideration), and are not the averaged values as in the energy confinement time database. The above five variables in Table 1 are related to plasma parameters, which are calculated from measured (and then fitted) profiles. The four variables below are related to the properties of magnetic configurations, and they are evaluated in modules for equilibrium in TASK3D-a (more precisely, VMEC [11] and GIOTA [12], as shown in Fig. 1 of Ref. [9]).

As a basis for grasping the fundamental feature of these variables, scatter plot matrix (after transferring variables to logarithmic scale for executing log-linear regression analyses afterwards) is shown in Fig. 3. The following features can be recognized from this figure.

(1) The range of log ρ_i* is small, which is mostly due to taking these values from experiments performed around 2.75 T only, targeting higher values of T_i. Addition of results from experiments performed at lower magnetic field strength certainly increases the size of the database. However, in such a case, the range of T_i becomes rather limited.

(2) The wedge-shaped distribution is recognized in log ρ_i* - log (T_e/T_i). The change of distribution seen at lower log ρ_i* values are coming from data close to the plasma edge (say, r_{eff}/a_{99}>0.9), where r_{eff} corresponds to the radius of the equivalent simple (i.e., circular cross section) torus in which the same volume is enclosed for the flux surface of interest, and a_{99} is the effective minor radius inside of which 99 % of the electron pressure is enclosed.

(3) The four variables (at right bottom corner) related to magnetic configurations have strong collinearity (in particular, among ϵ_{h}, ϵ_{t}, and ϵ_{eff}). This is mainly because the employed magnetic configuration for this database is fixed as that with vacuum magnetic axis position of 3.6 m for targeting higher values of T_{i}.

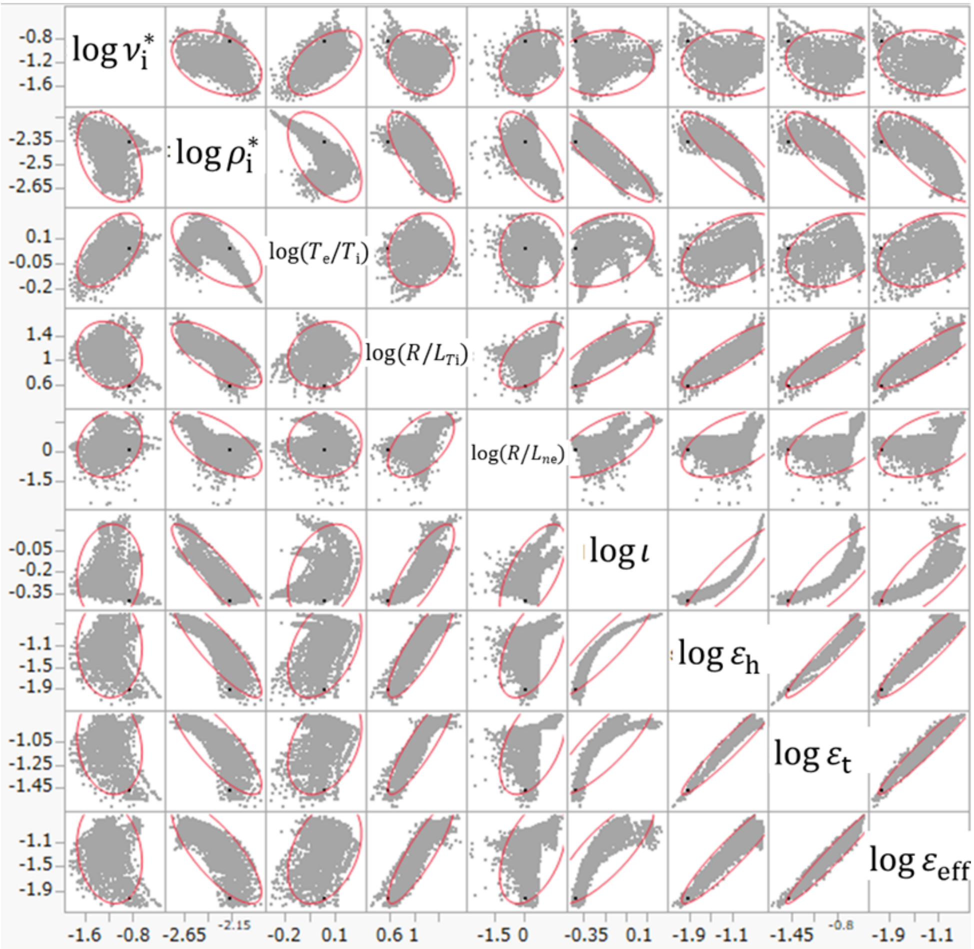

Fig. 3: Scatter plot matrix (after transferring variables to logarithmic scale for executing log-linear regression analyses afterwards) for all of nine variables listed in Table 1. The ellipses appearing in each panel are 95 % density ellipses, in which 95 % of data points are enclosed.

4. Model induction based on statistical approach

As attempted in Ref. [1], an inductive approach to acquire a simple expression for χ_i is also performed in this section. The different features than those of Ref. [1] are χ_i is normalized by $r_{\mathrm{eff}}^2\omega_{\mathrm{i}}$ (having dimension of m^2/s, with ω_{i} being the ion cyclotron frequency), and the AIC is considered as a measure for the reasonable model selection. Although the existence of a strong collinearity among some variables is known from Fig. 3, *a priori* selection of variables is not performed, and then all nine variables are kept considered to be subject to the AIC estimate. Fig. 4(a) shows the evolution of AICc value, which shows the step-shaped decrease according to the addition of variables. The details for the contribution of each variable will be described later. The finding from Fig. 4(a) is that goodness of the model using all nine variables is supported by AIC, although models using seven or eight variables give similar AICc value. Figure 4(b) shows the comparison of $\frac{\chi_i}{r_{\mathrm{eff}}^2 \omega_{\mathrm{i}}}$ values between database (vertical) and predicted (with 9 variables) values. The coefficient of determination, R^2, reaches as high as 0.95, with RMSE of 0.1941. These values are statistically promising, and better than those in Ref. [1]. The induced regression expression with all 9 variables is

$$\frac{\chi_i}{r_{\mathrm{eff}}^2 \omega_{\mathrm{i}}} = 10^{-27.9}\, \nu_{\mathrm{i}}^{*-0.65} \rho_{\mathrm{i}}^{*-3.42} (T_{\mathrm{e}}/T_{\mathrm{i}})^{1.26} (R/L_{T\mathrm{i}})^{-1.87} (R/L_{ne})^{0.033} (\iota/2\pi)^{1.27} \epsilon_{\mathrm{h}}^{-2.04} \epsilon_{\mathrm{t}}^{-4.56} \epsilon_{\mathrm{eff}}^{1.64} \text{ (2).}$$

It should be emphasized that this single line expression reproduces, with rather high value of R^2, as much as 3000 χ_i data which are covering core to the edge of plasmas included in the database. Thus, it can be powerfully used to reproduce or even predict T_{i} profiles in the plasma parameter regime covered by the database as shown in Fig. 2. Such systematic “validation” calculations should be demonstrated in the near future after this new way of inducing a thermal transport model is widely evaluated. When and if the fit to the database (or to the reproduction) is pursued with a high priority, such as after the extensive “learning” experiments, this approach that omits no available variables even with the existence of collinearity, is worthwhile as shown in Fig. 4(b). It should be emphasized that the extrapolation cannot be assured since this is just a regression fit to the database. Thus, expression (2) is not applicable beyond the employed database (Fig. 2), neither on other devices nor other operation scenario even in LHD. This approach needs “leaning” experiments for accumulating database to be regressed. The strong dependence of energy confinement time on magnetic field strength is also the out of scope, since the employed database is only on configurations with 2.75 T (cf., Sec. 3). However, expression (2) is a simple and good model for the employed database in LHD for the operation scenario targeting high-T_{i}.

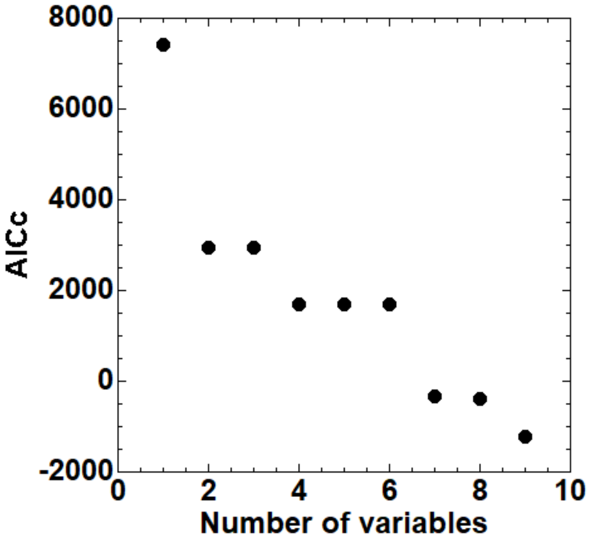

Fig. 4(a): The evolution of AICc value as a function of the number of variables for all nine variables listed in Table 1 and used in expression (2).

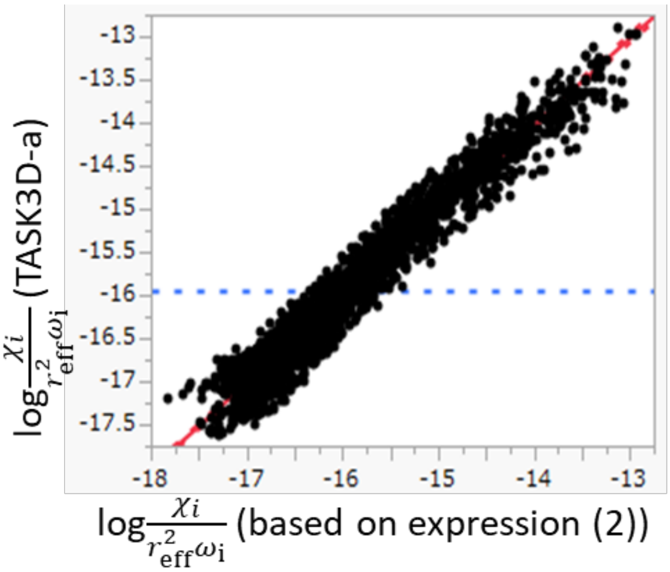

Fig. 4(b): The comparison of $\log(\chi_i/r_{\mathrm{eff}}^2\omega_{\mathrm{i}})$ values between TASK3D-a analysis database and the regression results based on expression (2).

The single line expression, expression (2), is easy to use for the reproduction of entire T_{i} profiles. However, when it comes to attempting a physics interpretation of the regression result, keeping a smaller number of variables having larger influence for the regression should be pursued. In this regard, an approach for pursuing this is explained below. As pointed out in Ref. [13], it would be worthwhile to divide the database radially, reflecting physics phenomenon such as the formation of ion internal transport barrier [14]. For this, the database for $r_{\rm eff}/a_{99}$<0.35 (404 data, roughly inside ion internal transport barrier) is utilized in the following discussion.

As shown in Fig. 1, the evolution of AICc value according to the addition of explanatory variables can be obtained through searching for the minimum value of AICc. The influence of variables onto the regression fit can be quantitatively measured by utilizing this information.

Figure 5(a) shows the comparison of $\frac{\chi_i}{r_{\rm eff}^2\omega_{\rm i}}$ values between database (vertical) and predicted (with all nine variables) values. This is the case for the minimum AICc with all nine variables kept. The coefficient of determination, R^2, reaches as high as 0.93, which is comparable to that in Fig. 4(b). The evolution of AICc values are shown in Fig. 5(b). The first three points correspond to $\nu_{\rm i}$*, $\rho_{\rm i}$* and $T_{\rm e}/T_{\rm i}$, as *a priori* utilized in Ref. [1]. Along the way reaching the AICc minimum state, two obvious sudden decreases of AICc are found. They correspond to the inclusion of $R/L_{\rm Ti}$ (4th point) and $\epsilon_{\rm h}$ (7th point, weaker influence than $R/L_{\rm Ti}$ though), both of which should be kept in the regression. Since a strong collinearity (0.835) exists between $\nu_{\rm i}$* and $\rho_{\rm i}$*, as shown in Fig. 5(c), one of these two variables ($\rho_{\rm i}$* in this case) is kept. Thus, the regression with 4 variables, ($\rho_{\rm i}$*, $T_{\rm e}/T_{\rm i}$, $R/L_{\rm Ti}$ and $\epsilon_{\rm h}$) is performed by K=4 cross validation in which dataset is divided into four groups, and three out of four groups are subsequently used for regression, and the remaining one group is used for its validation. Then the obtained four models are averaged. Result is shown in Fig. 5(d), worsening the goodness of a model a bit than that for a case with all nine variables, as indicated by R^2 of 0.87 and RMSE of 0.188. However, R^2 is keeping a practically high value. Table 2 summarizes goodness of models by varying the total number of variables (9, 4 and 3 with 4 variations). For cases with 3 variables, it is clearly shown that excluding $R/L_{\rm Ti}$ significantly worsens the goodness, as is expected from AICc consideration (Fig. 5(b)). If $R/L_{\rm Ti}$ is kept in the regression with 3 variables, any other combinations of 2 variables give similar goodness values (somewhat better if $\rho_{\rm i}$*is excluded, possibly due to the narrow range of $\rho_{\rm i}$*in the current database). In such a way, several regression results could be considered for their relevance and implications for physical interpretation.

Here, the regression expression for the case of 4 variables,

$$\frac{\chi_i}{r_{\rm eff}^2\omega_{\rm i}} = 10^{-0.15}\,\rho_{\rm i}^{*6.42}(T_{\rm e}/T_{\rm i})^{5.35}(R/L_{T{\rm i}})^{-2.16}\epsilon_{\rm h}^{-1.35} \quad (3),$$

is selected for discussing possible physics interpretations. The negative power of $R/L_{\rm Ti}$ would be physically interpreted as that the ion heat diffusivity in ion internal transport barrier decreases as the ion temperature gradient becomes larger (or its scale length becomes smaller making $R/L_{\rm Ti}$ larger) [14]. The increase of heat diffusivity (then the decrease of the ion temperature gradient) for increasing $T_{\rm e}/T_{\rm i}$ has also been recognized experimentally in LHD [14]. Large value of power of $\rho_{\rm i}$* is considered to indicate that the ion heat diffusivity becomes larger as the ion temperature becomes higher. The $\epsilon_{\rm h}$ itself currently does not cover any of the changes of magnetic configurations (cf., description (3) in

Sec. 3). However, its exclusion worsens the goodness as seen in Table 2. Based on these considerations, expression (3) statistically infers complicated balance between ion temperature (and its ratio to electron temperature, as well) and its gradient and the feature of transport barrier formation, in a single line expression. It would be interesting to check this kind of regression results could be reasonable against large-scale cutting-edge simulations when they accumulate plenty of simulation results.

It should be mentioned here that it has not been possible to obtain satisfactory goodness of fits with limited number (eg., 4 for r_{eff}/a_{99}<0.35 as described above) of explanatory variables for outer radial regions such as 0.4<r_{eff}/a_{99}<0.6 and 0.7<r_{eff}/a_{99}<0.9. This is clearly due to the rather narrow range of T_{i} in the employed database for those regions. Extending database would resolve this, and the similar deduction of physics interpretations will be tried then.

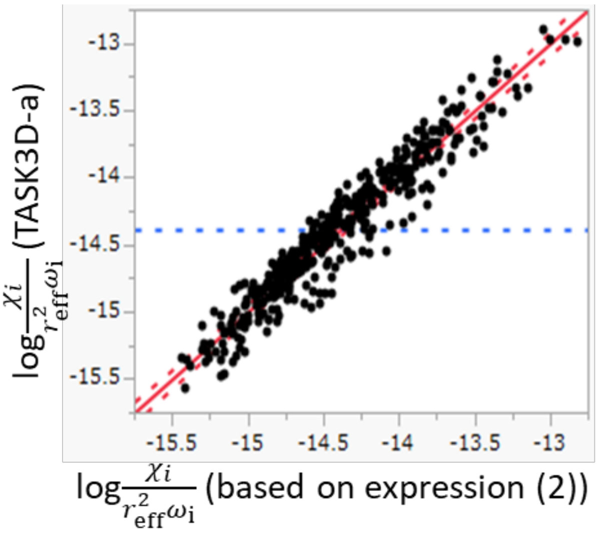

Fig. 5(a): The comparison of $\log(\chi_i/r_{\mathrm{eff}}^2\omega_{\mathrm{i}})$ values between TASK3D-a analysis database and the regression results based on expression (2) for database for r_{eff}/a_{99}<0.35.

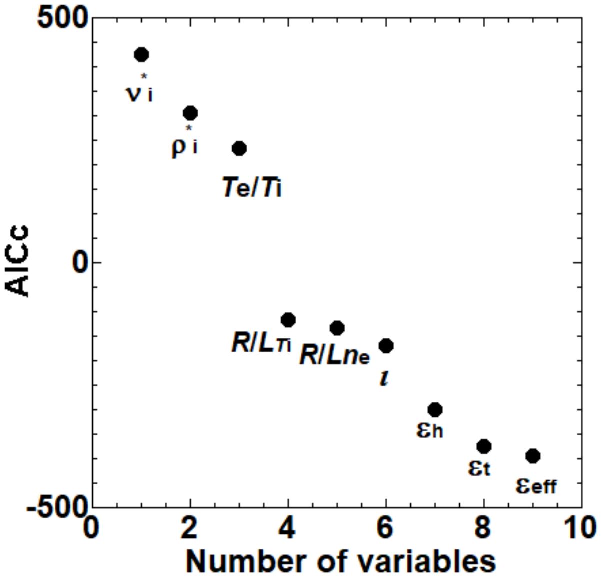

Fig. 5(b): The evolution of AICc value as a function of the number of variables for all nine variables listed in Table 1 and used in expression (2).

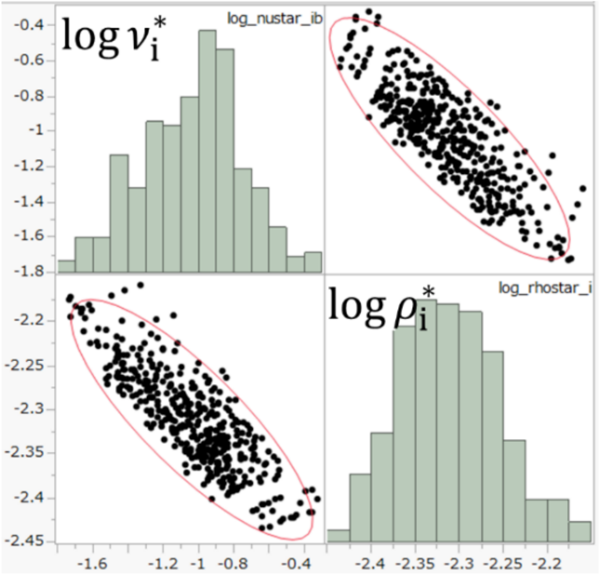

Fig. 5(c): Scatter plot matrix (after transferring variables to logarithmic scale for executing log-linear regression analyses afterwards) for ν_i* and ρ_i* for database for r_{eff}/a_{99}<0.35.

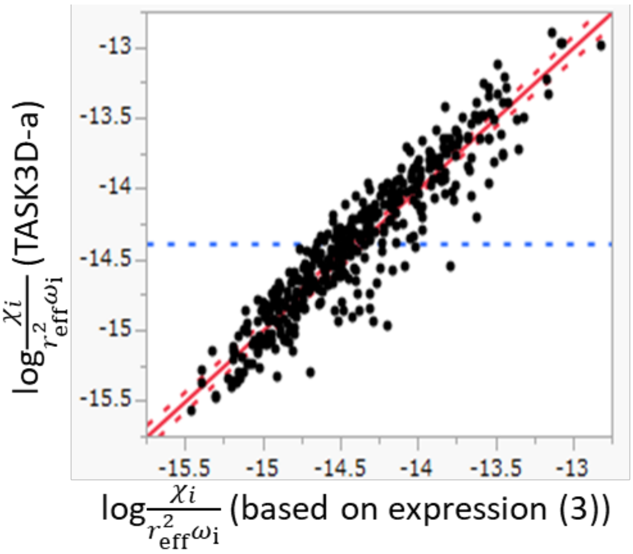

Fig. 5(d): The comparison of $\log(\chi_i/r_{\rm eff}^2\omega_{\rm i})$ values between TASK3D-a analysis database and the regression results based on expression (3) for database for $r_{\rm eff}/a_{99}$<0.35.

Number of variables	***R*²**	**RMSE**	**AICc**
9	0.93	0.146	-394.1
4 ($\rho_{\rm i}$*, $T_{\rm e}/T_{\rm i}$, $R/L_{\rm Ti}$, $\epsilon_{\rm h}$)	0.87	0.188	-197.3
3 ($\rho_{\rm i}$*, $R/L_{\rm Ti}$, $\epsilon_{\rm h}$)	0.8	0.234	-19.6
3 ($\rho_{\rm i}$*, $T_{\rm e}/T_{\rm i}$, $R/L_{\rm Ti}$)	0.81	0.234	-19.6
3 ($T_{\rm e}/T_{\rm i}$, $R/L_{\rm Ti}$, $\epsilon_{\rm h}$)	0.84	0.213	-98.9
3 ($\rho_{\rm i}$*, $T_{\rm e}/T_{\rm i}$, $\epsilon_{\rm h}$)	0.69	0.296	168.6

Table 2: Summary of goodness of models (R^2, RMSE, and AICc) for varying the total number of variables (nine, four and three with four combinations) for database for $r_{\rm eff}/a_{99}$<0.35.

5. Summary and discussion

The statistical approach for inducing a transport model (for the ion heat diffusivity, as an example) has been progressed based on its initial publication [1] in two folds. One is the reproducibility of the values themselves, and the other is the link to the physics interpretation through selection of variables by means of application of AIC.

This approach, by nature, provides practical description of ion heat diffusivity over the parameter range which is covered by analyses database, and does not guarantee the predictability beyond such a parameter range. In other words, this approach is for reproduction purposes, based on "learning" experiments. However, it is indeed rather simple, and this approach itself (not the obtained expressions in this Letter) could be used as guidance for the real-time operation of current devices and even of future fusion reactors, if it could provide learning experiments in its initial phase of operation. Moreover, the induced parameter dependence, which are power to the selected variables, could establish a link to the rigorous cutting-edge transport simulations, in terms of providing a guidance for expected parameter dependence to be reproduced by them.

In the meantime, the database has been extended through the extensive applications of the integrated transport analysis suite, TASK3D-a, to the progressing LHD experiments. The covered parameter ranges will be expanded accordingly, which would provide further opportunity to extend this approach, for example, for the electron heat transport and for the particle transport.

Acknowledgements

The author acknowledges the Collaboration Research Startup Initiative of the Institute of Statistical Mathematics (ISM). The TASK3D-UD (Users and Developers) and the LHD Experiment Group are also appreciated for formulating the TASK3D-a transport analyses database, which is an essential data source for this study. This work has been supported by the NIFS Collaborative Research Programs, NIFS14KNTT025 and NIFS17UNTT008, ISM Cooperative Research Program (2017-ISMCRP-1-1028, 2018-ISMCRP-1-1002 and 2019-ISMCRP-2027), and NINS (National Institutes of Natural Sciences) program for cross-disciplinary study (Grant Numbers 01321802 and 01311904) on Turbulence, Transport, and Heating Dynamics in Laboratory and Solar/Astrophysical Plasmas: "SoLaBo-X".

References

[1] M. Yokoyama et al., Plasma Fusion Res. **9** (2014) 1302137.

[2] O. Meneghini et al., Phys. Plasmas **21** (2014) 060702.

[3] E. Narita et al., Plasma Phys. Control. Fusion **60** (2018) 025027.

[4] H. Akaike, "Information theory and an extension of the maximum likelihood principle", Proceedings of the 2nd International Symposium on Information Theory, Petrov, B. N., and Caski, F. (eds.), Akadimiai Kiado, Budapest (1973) 267.

[5] N. Sugiura, Communications in Statistics - Theory and Methods **7** (1978) 13.

[6] G. Schwarz, The Annals of Statistics, **6** (1978) 461.

[7] H. Yamada et al., Nucl. Fusion **45** (2005) 1684.

[8] Y. Takeiri et al., Nucl. Fusion **57** (2017) 102023.

[9] M. Yokoyama et al., Nucl. Fusion **57** (2017) 126016.

[10] C. Suzuki et al., Plasma Phys. Control. Fusion 55 (2013) 014016.

[11] S.P. Hirshman and J.C. Whiston, Phys. Fluids **26** (1983) 3553.

[12] M. Yokoyama et al., Research Report NIFS-810, National Institute for Fusion Science, Japan (2005) on the numerical code originally developed by L. Hedrick (retired, Oak Ridge National

Laboratory).

[13] M. Yokoyama and H. Yamaguchi, Plasma Fusion Res. **14** (2019) 1303095.

[14] K. Nagaoka et al., Nucl. Fusion **55** (2015) 113020.