*31st EPS Conference on Plasma Phys. London, 28 June - 2 July 2004 ECA Vol.**28G**, O-2.03 (2004)*

Integrated Interpretative Transport Modelling

A. Dinklage[1], C. D. Beidler[1], R. Fischer[2], H. Maaßberg[1], J. Svensson[1], Yu. A. Turkin[1]

Max-Planck-Institut für Plasmaphysik, EURATOM Association, [1]Greifswald, [2]Garching, Germany

Integrated analysis of fusion data

Usually, diagnostic data and modelling results are dealt with separately in the analysis of fusion data, i.e., experimental results are passed to modelling codes for addressing physical issues. But interdependencies of modelled and measured parameters are numerous. In addition to the increase of the reliability of physics results, these entanglements allow for accessing barely measurable quantities more accurately, e.g., particle and heat fluxes. The goal of this paper is to introduce a framework for inclusion of modelling results to data analysis. Application to W7-AS data and transport modelling will be discussed.

The notion of *integrated modelling* and *integrated data analysis* [1] is used here for the concise combination of different sources of information such as different diagnostics results. The methods of Bayesian probability theory will be employed in order to arrive at a comprehensive modelling and data analysis. Technically, the approach uses raw data and their respective uncertainties and yields estimates for all parameters entering the model, in particular those quantities, which are of physical interest. The motivation for a framework allowing for integrated analyses arises from physically motivated requirements in large, long-pulse fusion devices. E.g., the operation of W7-X plasmas on time scales large with respect to the typical configuration relaxation time ($\tau_{L/R} > 10\,\mathrm{s}$) requires formalized on-line analyses of validated physics quantities. Moreover, huge amounts of data require automated approaches. For detailed off-line investigations, the inclusion of interdependencies enhances both the reliability of analyses and evidence for quantities which are not directly measurable.

A simple picture of the integrated approach consists of modules representing individual diagnostics measurements or modelling results, all of which are linked through physical relationships. An ubiquitous case for such links in fusion experiments is mapping of different, spatially resolved measurements onto a common grid given by magnetic equilibria. The example discussed in this paper focuses on transport models. Probabilistic methods are briefly discussed as a prerequisite of integrated modelling. In addition to several benefits to be discussed in this paper, Bayesian probability theory also allows one to deal with so-called systematic errors, which is essential for a framework for joint treatment of different diagnostics and modelling. Aspects of software architecture, such as WebServices or GRID computing, allowing the communication of codes are beneficial for the realization of the integrated approach; more details can be found in Ref. [2].

The Bayesian approach

Bayesian probability theory provides a framework for inference from uncertain data; the uncertainties are encoded as probability density functions (PDF) . The outcome is the conditional PDF of all model parameters (including the quantities of interest) *given* some data. In the Bayesian approach PDFs are used to quantify uncertain information rather than frequency distributions of random variables. For the purpose of inference, all sources of uncertainties are to be quantified. Then, the outcome, the so-called *posterior* PDF is given by Bayes' theorem multiplying the *likelihood* PDF (uncertainties of data) and the *a priori* knowledge (without data). The *prior* PDF may encode physics knowledge (e.g. temperatures cannot be negative) or may be derived from information criteria [3].

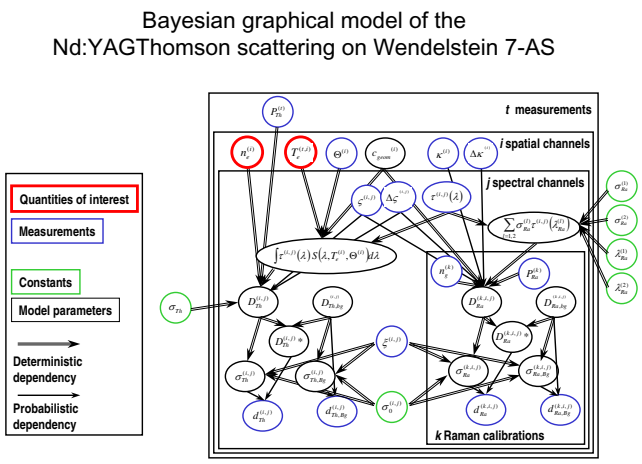

Figure 1: Bayesian graphical model for the Nd:YAG Thomson scattering on W7-AS.

The Bayesian approach treats quantities of interest and nuisance parameters similarly – the latter parameters affect the result and introduce additional uncertainties (e.g. calibration factors). The classification of parameters to be of nuisance or of interest depends on the application and physical issue to be addressed. It is to be noted that consideration of nuisance parameters may give valuable information for the optimization of fusion diagnostics [4].

From a formal point of view the integration of different diagnostics to *diagnostic sets* is straightforward. An advantageous semantics is the formulation of interdependencies within Bayesian Graphical Models [5], which is a convenient basis for the specification of the full posterior PDF. These *directed acyclic graphs* can be used to built up an appropriate representation of a large, complex probabilistic model, being an appropriate framework for the entanglements typically occurring in fusion data analysis. The Bayesian graphical model may also reflect the topology of different software modules to be linked. The outcome of the graphical model is the joint PDF of the set of all nodes U ($U = \{d, p\}$ comprises both the data d and model parameters p) which can be shown to be the product of the conditional distribution of each node given their parents, i.e., $p(U) = \prod_{u \in U} p\left(u|\mathrm{parents}[u]\right)$ [2].

As an example for a realistic probabilistic model, Fig. 1 shows a Bayesian graphical model for the Nd:YAG Thomson scattering system operated on Wendelstein 7-AS, adopted from a similar statistical model in Ref. [4].

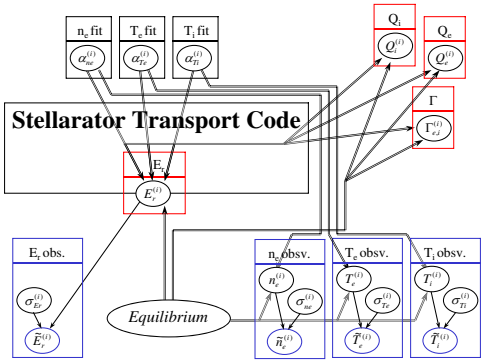

Figure 2: Simplified Bayesian graphical model for an integrated interpretative modelling employing the ST code [6].

In a physics-motivated approach the quantities of interest (for clarity red colored in Fig. 1) represent the canonical links of different diagnostics modules. A second graphical model shown in Fig. 2 depicts the links of different diagnostics results (which enter in Fig. 2 as data with errors). The linkage is given by the interpretative module of the stellarator transport code ST [6]. Uncertainties of any node, i.e. any quantity or parameter in the graph, can be derived. As a different example, a Bayesian graphical model for the diagnostics of Wendelstein 7-AS was used to determine the uncertainty of mapping and ι profiles.

Results for Wendelstein 7-AS

Figure 3 depicts results derived with the graphical model shown in Fig. 2. These results represent marginal posterior distributions from a Markov-Chain-Monte-Carlo sampling of the full posterior PDF. That PDF P_p is given coarsely a χ^2-like expression and the discrepancy of the entire model is measured by a weighted sum of deviations:
$\ln P_p = -\sum_{i,j,..}\left((n_e^{(i)} - f_{ne}^{(i)}(r_{eff}))/\sigma_{ne}^{(i)}\right)^2 + .. + \left((E_r^{(j)} - g_{Er}^{(j)}(r_{eff}, f_{ne}^{(j)}, f_{Te}^{(j)}, f_{Ti}^{(j)}))/\sigma_{Er}^{(i)}\right)^2$.
It is the functional dependency g which introduces the link between density and temperatures on the one hand and the radial electric field on the other. Practically, g is provided by the transport code module. The implicit dependence between radial electric field E_r and density and temperature profiles is obtained from neoclassical transport. In order to check the validity of the transport model, experimental fluxes derived from heat and particle sources are shown in Fig. 3 as well. The radial electric field is determined from the ambipolarity condition.
The shade of the patches close to lines in Fig. 3 is a measure of uncertainty of the integrated model for experimental data from a W7-AS high ion temperature shot given.
Physically, it can be concluded that particle and energy transport are quite consistent with neoclassical approximations within the core of a high ion temperature plasma of Wendelstein 7-AS. The edge fluxes, however, considerably differ from neoclassical predictions and anomalous transport dominates. The radial electric field at the edge, however, is well described by

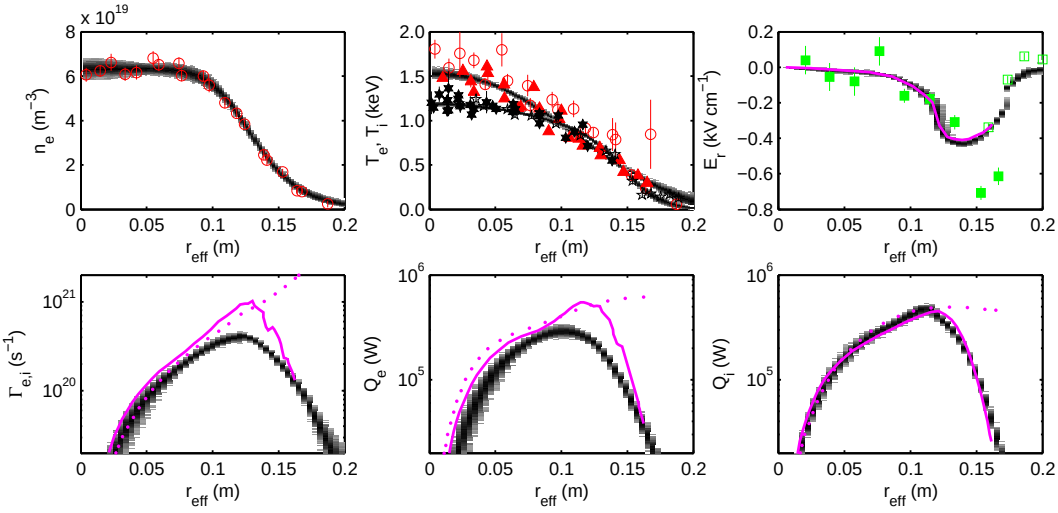

Figure 3: Results for integrated transport modelling according to the graphical model shown in Fig. 2 for a W7-AS high-T_i shot. Magenta dotted lines result from integration of sources, magenta solid lines are the result from sequential analysis. The error of the integrated modelling are depicted by the gray patches, the scale of which is proportional to the logarithm of the PDF and black corresponds to maximum probability.

neoclassical theory. Hence, the anomalous contributions at the plasma edge of W7-AS high ion temperature discharges appears to be intrinsically ambipolar. A possible mechanism for such an anomalous edge transport in the W7-AS shot analyzed is electrostatic turbulence.

Concluding, methods for an interpretative integrated modelling were demonstrated and elementary steps, such as the probabilistic modelling of single diagnostics and the linkage of different diagnostics were discussed. The integration makes use of the Bayesian probability theory. Applicability of methods is tested on W7-AS data in order to prepare tools for Wendelstein 7-X, which can be automated (for on-line analyses) and allow one to determine barely measurable quantities and to quantify the reliability of physical results. The Bayesian framework allows one to include physics results at different stages of complexity of data analysis.

References

[1] R. Fischer, A. Dinklage , E. Pasch, *Plasma Phys. Contr. Fusion* **45**, 1095 (2003).

[2] J. Svensson, A. Dinklage, J. Geiger, A. Werner, R. Fischer, *Rev. Sci. Instrum.* (at press 2004).

[3] A possibility is the *reference prior* idea. See: J.M. Bernardo, *J. Statist. Pl. Inf.* **65**, 159 (1997).

[4] R.Fischer, C.Wendland, A.Dinklage, S.Gori, V.Dose, *Plasma Phys. Contr. Fusion* **44**, 1501 (2002).

[5] S.L. Lauritzen, A. P. Dawid, B. N. Larsen, H.-G. Leimer, *Networks* **20**, 491 (1990); see also D. J. Spiegelhalter, N. G. Best, W. R. Gilks, H. Inskip, in *Markov Chain Monte Carlo in Practice*, ed. W. R. Gilks, S. Richardson, D. J. Spiegelhalter, (Chapman & Hall, London, 1996), pp. 21-43.

[6] Y. Turkin et al., this conference (P1-198).