ELSEVIER

Journal of Nuclear Materials 288 (2001) 153–162

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Evaluation of chemical erosion data for carbon materials at high ion fluxes using Bayesian probability theory

V. Dose *, R. Preuss, J. Roth

Centre for Inter-disciplinary Plasma Science, Max-Planck-Institut für Plasmaphysik, EURATOM Association, POB 1322, D-85741 Garching bei München, Germany

Received 3 September 2000; accepted 26 October 2000

Abstract

Parameters for a function modeling chemical erosion of carbon as a function of particle energy and flux as well as target temperature are determined from the available data. The analysis is based on Bayesian inference and provides posterior estimates of parameters, erosion yields and associated error margins. The model function incorporates the flux dependence via a time constant for hydration. The present analysis yields a hydration time of 0.22 m s independent of surface temperature and isotope mass.

1. Introduction

Chemical erosion due to hydrogen ion bombardment is the dominant erosion process for carbon-based plasma facing materials in thermonuclear fusion devices [1]. In today's fusion experiments in general large areas of the main vessel walls are covered by carbon tiles and the areas of direct plasma-wall contact, such as divertor plates or limiters, are made from carbon-fiber composite (CFC) material [2–4]. While the vessel walls are subject to low flux neutral irradiation ($\sim 10^{20}/\mathrm{m}^2$ s) and remain at ambient temperature, the divertor plates experience ion fluxes up to $10^{24}/\mathrm{m}^2$ s and, consequently, elevated surface temperatures will occur [5]. In both cases, the ion energy distribution will be broad, with highest fluxes in the 10–50 eV range. In spite of the low fluxes the impurity content of the plasma will be dominated by wall erosion due to the large area of the walls. The estimated erosion at the divertor plates leads to severe lifetime problems of the components. Both erosion processes result in the deposition of carbon layers within the machine retaining unacceptably high inventories of hydrogen isotopes [6].

In the low flux regime, i.e., below $10^{19}/\mathrm{m}^2$ s, the mechanism of chemical erosion is reasonably well understood. From ion beam experiments [7] and experiments using thermal hydrogen atoms [8] it could be established that the total chemical erosion results from different basic processes: within the ion range carbon atoms with unsaturated bonds will be fully hydrogenated already at room temperature. The concentration of available bonds, C_{H}, may be enhanced by radiation damage processes if the ion energy exceeds about 15 eV. As the temperature is increased thermal bond-breaking processes become possible and CH_3 or C_2H_2 radicals are released resulting in the thermal erosion process, Y_{therm}. A further increase in temperature enables hydrogen loss due to hydrogen molecule formation in Eley–Rideal processes and chemical erosion decreases. Thus, a maximum of chemical erosion, Y_{therm}, is established at intermediate temperatures ranging from 500 to 900 K depending on ion flux. At room temperature no thermal release of hydrocarbon radicals is possible. However, these radicals are bound to the surface with a much lower binding energy than non-hydrogenated carbon atoms. Consequently, surface radicals can be sputtered already at ion energies well below the threshold energy for physical sputtering of carbon and from the available data a threshold energy of 2 eV can be estimated for chemically enhanced physical sputtering, Y_{surf}. These processes have been described analytically and formulae for C_{H}, Y_{therm} and Y_{surf} are given in [7].

At high fluxes Φ, such as experienced in fusion devices, there was indication from scattered data that the

* Corresponding author.

0022-3115/01/$ - see front matter
PII: S0022-3115(00)00709-1

chemical sputtering yield decreases with ion flux [9]. The data base did not allow, however, one to draw quantitative conclusions on such a flux dependence [10]. On the other hand, deposition experiments for hydrogenated carbon layers indicated that hydrogenation may include steps requiring times of the order of ms, thus limiting the hydrogen content of films if the deposition rate gets too fast [11]. Such a time dependent step would consequently reduce the concentration of hydrocarbon radicals in the surface, limit the eroded molecule flux to the flux independent rate at which the radicals can be formed and lead to a $1/\Phi$ dependence of the chemical erosion yield at high fluxes. The analytic description in [7] includes this assumption of a flux dependence.

Within the past two years new experimental efforts have been made to investigate the flux dependence, both at room temperature and around the maximum temperature for chemical erosion, T_{max}. In plasma simulators [12,13] ion fluxes of the order of $10^{23}/m^2$ s are reached and new data are available from dedicated experiments in the divertors of tokamaks [14,15]. In most of these measurements an estimate of the error of the resulting erosion yield was given including the determination of the incident ion flux, of the spectroscopic determination of the CD band intensity and of the calculation of the emitted hydrocarbon flux from the measured CD band. In the present paper, these new data are evaluated on the basis of the analytic model including the plausible assumption of the yield decrease at high fluxes. The functional form of the model [7] is retained and fitted to the data with only two free parameters: a constant c, effectively determined by the low flux data from ion beam experiments and the parameter λ, determining the flux where the strong flux dependence sets in. Bayesian probability theory is employed in order to determine these parameters. From the bulk of the published low flux data the weight loss results have been selected [16] as the most simple and reliable total yields, since these data require no further calibration factors. Such factors are necessary for mass spectroscopy and optical emission spectroscopy. Ion energies near 30 eV are selected as typical for limiter application.

2. Bayesian rules

This section is intended to fix the nomenclature and definitions. For a more detailed introduction to Bayesian probability theory we refer interested readers to excellent papers in the literature which cover this subject [17–20].

Bayesian inference rests on the application of two rules. The first is the product rule which allows to expand the probability density function $P(H,D|I)$ depending on the two variables H and D conditional on further information I into two simpler factors

$$P(H,D|I) = P(H|I)P(D|H,I). \tag{1}$$

The first factor $P(H|I)$ depends now only on the variable H conditional on the previous background information I. The second factor depends also on a single variable D but the condition has been extended from a knowledge of I to both H and I. Let us identify H with hypotheses (physics) and D with data. Eq. (1) allows of course the alternative expansion due to symmetry in the variables on the left-hand side

$$P(H,D|I) = P(D|I)P(H|D,I). \tag{2}$$

Combination of (1) and (2) yields Bayes' theorem

$$P(H|D,I) = \frac{P(H|I)}{P(D|I)}P(D|H,I). \tag{3}$$

Bayes' theorem represents a recipe for learning. It tells us how to update our knowledge of physics $P(H|I)$ in the light of data obtained from a suitably designed experiment $P(D|H,I)$ into posterior knowledge $P(H|D,I)$. $P(H|I)$ is called the prior on the hypothesis H. $P(D|H,I)$ constitutes the theory of the experiment, e.g., the prediction of the data to be measured assuming the physics H to be known. This probability density is called the sampling distribution of the data when regarded as a function of D and is normalized in these variables, i.e., it is the predictive distribution for possible data sets given the same hypotheses H. Alternatively it is called the likelihood when considered as a function of H and the data are held constant. $P(H|D,I)$ is called the posterior probability distribution of H which combines the knowledge available prior to taking the data D with the result of the experiment. $P(D|I)$ finally is called the evidence or the prior predictive value or the global likelihood of the entire class of hypotheses $\{H\}$. We prefer the latter term because $P(D|I)$ is not independent of the other probabilities by virtue of the second rule of Bayesian probability theory, the sum rule. It states that

$$\begin{aligned} P(D|I) &= \int P(H,D|I)\,\mathrm{d}H \\ &= \int P(H|I)P(D|H,I)\,\mathrm{d}H. \end{aligned} \tag{4}$$

The global likelihood $P(D|I)$ results after (4) from an integration over all possibilities of H. In that case H is characterized by a set of parameters the symbolic integration over H means in practice integration over parameter space. $P(D|I)$ plays a crucial role in the comparison of different models given a data set D. Let us specify the background information I further by models M_1 or M_2 with associated different parameter sets H_1 and H_2. Bayesian probability theory allows one to answer the important question whether and to what degree model M_1 is to be preferred over model M_2 in the light of the experimental data D. To this end we calculate

$P(\mathrm{M}_i|D,I)$, which by application of Bayes' theorem is given by

$$P(\mathrm{M}_i|D,I) = \frac{P(\mathrm{M}_i|I)P(D|\mathrm{M}_i,I)}{P(D|I)}. \tag{5}$$

If we take the ratio of (5) for $i = 1$ and $i = 2$ we obtain the 'odds ratio'

$$\frac{P(\mathrm{M}_1|D,I)}{P(\mathrm{M}_2|D,I)} = \frac{P(\mathrm{M}_1|I)}{P(\mathrm{M}_2|I)}\frac{P(D|\mathrm{M}_1,I)}{P(D|\mathrm{M}_2,I)}. \tag{6}$$

The first factor on the right-hand side is called the prior odds. It specifies via a real number the experts' belief about how much model M_1 is to be preferred over model M_2. If this number deviates very much from one there is usually very little motivation for carrying out an experiment. At least it will be difficult to find funding for it! So most of all experimental work relates to the case that we have no reason to assign different values to $P(\mathrm{M}_1|I)$ and $P(\mathrm{M}_2|I)$. The odds ratio is then entirely determined by the second, the so-called Bayes factor. The required probability densities to calculate this factor are obtained from a specification of the parameters H_i entering model M_i by application of the sum rule

$$P(D|\mathrm{M}_i,I) = \int P(D,H_i|\mathrm{M}_i,I)\,\mathrm{d}H_i. \tag{7}$$

Further mathematical manipulations of probabilities in the rest of this paper are based on the employment of the basic rules (1) and (4).

3. The likelihood

Consider the case where two sets of data $\vec{\delta}$ and $\vec{\Delta}$ of measured erosion yields are available, were taken at fluxes $\vec{x}$ and X, and have standard deviation $\vec{\sigma}$ and Σ. The vector symbol comprises the whole data set, i.e., $\vec{\delta} = \{\delta_1, \ldots, \delta_{N_\delta}\}$, where $N_\delta(N_\Delta)$ is the number of the data $\vec{\delta}(\vec{\Delta})$. The usual situation in experimental physics is that $\vec{x}$ and $\vec{X}$ cover different ranges with everything between no overlap and full overlap and that $\vec{\delta}$ and $\vec{\Delta}$ suffer from different calibrations. Our problem is then to determine model parameters on the basis of all available data. As already mentioned in the introduction the model function is fully specified by a linear parameter c establishing the scale of the fit function and another non-linear parameter λ related to the flux at which the erosion yield starts to decrease as a function of flux. Let us further introduce a calibration parameter γ which, when applied to the data set $\vec{\Delta}$, rescales it on the same scale as $\vec{\delta}$. All this information can be condensed in the likelihood

$$P\left(\vec{\delta},\vec{\Delta}\middle|\vec{x},\vec{X},\vec{\sigma},\vec{\Sigma},c,\lambda,\gamma,\mathrm{M},I\right) = P\left(\vec{\delta}\middle|\vec{x},\vec{X},\vec{\sigma},\vec{\Sigma},c,\lambda,\gamma,I\right)P\left(\vec{\Delta}\middle|\vec{\delta},\vec{x},\vec{X},\vec{\sigma},\vec{\Sigma},c,\lambda,\gamma,\mathrm{M},I\right). \tag{8}$$

The right-hand side of (8) follows from the product rule and may be further simplified. Of course the probability of the data set $\vec{\delta}$ is independent of the flux $\vec{X}$ used to collect the data $\vec{\Delta}$ in another experiment. By the same token, $P(\vec{\Delta}|\vec{\delta},\vec{x},\vec{X},\vec{\sigma},\vec{\Sigma},c,\lambda,\gamma,\mathrm{M},I)$ is independent of $\vec{x}$, $\vec{\delta}$ and $\vec{\sigma}$ for a truly independent experiment. Moreover, we have chosen the convention that the data set $\vec{\delta}$ does not require recalibration and therefore $P(\vec{\delta}|\vec{x},\vec{X},\vec{\sigma},\vec{\Sigma},c,\lambda,\gamma,I)$ is also independent of γ. In the following we shall simplify our notation by simply suppressing conditioning variables which make no sense on the grounds of logical independence. The simplified Eq. (8) is then

$$P\left(\vec{\delta},\vec{\Delta}\middle|\vec{x},\vec{X},\vec{\sigma},\vec{\Sigma},c,\lambda,\gamma,\mathrm{M},I\right) = P\left(\vec{\delta}\middle|\vec{x},\vec{\sigma},c,\lambda,I\right)P\left(\vec{\Delta}\middle|\vec{X},\vec{\Sigma},c,\lambda,\gamma,\mathrm{M},I\right). \tag{9}$$

We now model our data set $\vec{\delta}$ as

$$\delta_i = c\varphi(x_i,\lambda) + \alpha_i, \tag{10}$$

where α_i denotes the error associated with the measurement of δ_i and the expectation value $\langle\alpha_i\rangle$ of α_i is of course zero. We assume further, that the variance of the true error α_i is $\langle\alpha_i^2\rangle = \sigma_i^2$. This allows us to express the first factor in (9) by application of the principle of maximum entropy as [20]

$$P\left(\vec{\delta}\middle|\vec{x},\vec{\sigma},c,\lambda,I\right) = \prod_{i=1}^{N_\delta}\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left\{-\frac{1}{2}[\delta_i - c\varphi(x_i,\lambda)]\Big/\sigma_i^2\right\}. \tag{11}$$

For the second factor in (9) we shall consider two different cases. Let us first assume that the optical system is absolutely calibrated for both the hydrogen and the CH band. In this case the conversion of the hydrogen spectroscopic data to flux values is free of error due to the favorable atomic data situation for hydrogen, while cross-section data and optical transition probabilities for CH_x are less well-known and the conversion of CH band emissions to erosion data may suffer from a bias which we account for by the factor γ (model M_1). We then have

$$\gamma\Delta_j = c\varphi(X_j,\lambda) + \beta_j \tag{12}$$

and consequently assuming $\langle\beta_j\rangle = 0$ and $\langle\beta_j^2\rangle = \Sigma_j^2$ the likelihood function becomes

$$P\left(\vec{\Delta}\middle|\vec{X},\vec{\Sigma},c,\lambda,\gamma,I\right) = \prod_{j=1}^{N_A}\frac{\gamma}{\Sigma_j\sqrt{2\pi}}\exp\left\{-\frac{1}{2}[\gamma\Delta_j - c\varphi(X_j,\lambda)]^2\Big/\Sigma_j^2\right\}. \tag{13}$$

However, with erosion data collected in fusion machines there is another conceivable situation. The optical system used to record hydrogen and CH band emissions may suffer from a calibration error which translates into a common recalibration factor γ for both the hydrogen flux X_j and the erosion yield Δ_j (model M_2). In this case we would have to model our data according to

$$\gamma\Delta_j = c\varphi(\gamma X_j,\lambda) + \beta_j. \tag{14}$$

The two formulations (12) and (14) for the data set $\vec{\Delta}$ constitute the two models M_1 and M_2 which we shall compare by calculating the odds ratio (6).

We have nearly, but not quite, completed the specification of the likelihood function. Both (11) and (13) require the knowledge of the true errors σ_i and Σ_i, respectively. Except from numerical simulations these are hardly ever known. What we may know (and do know in the present case) are experimental estimates of the true errors σ_i and Σ_i which we denote by s_i and S_i, respectively. These estimates may come close to the true values as for example in counting experiments where Poisson statistics apply. Quite often, however, we encounter the case that the quoted experimental error is either obviously too large, or the scatter of equivalent data is much larger than expected from the size of the quoted error. We usually talk about outliers in the latter case. They are particularly harmful since an outlier with a small specified estimate of the true error will strongly bias the conclusions drawn from the data set. Outlier tolerance in our inferences may be obtained in the following way [21]. Let us assume that the probability density for the true error σ is given in terms of two parameters μ and v by the distribution

$$P(\sigma_i|s_i,\mu,v,I) = 2\frac{\mu^v}{\Gamma(v)}\left(\frac{s_i}{\sigma_i}\right)^{2v}\mathrm{e}^{-\mu s_i^2/\sigma_i^2}\frac{1}{\sigma_i}. \tag{15}$$

The solid curve in Fig. 1 displays (15) as a function of σ/s for values of the conditioning parameters μ, v which we shall fix now. Our first obvious requirement is that $\langle\sigma\rangle = s$. That means that we assume that the experimenter has tried hard and supplied a bias free error estimate. The expectation value $\langle\sigma\rangle$ is then given in terms of μ and v

$$\frac{\langle\sigma\rangle}{s} = \sqrt{\mu}\frac{\Gamma(v-1/2)}{\Gamma(v)} \stackrel{!}{=} 1. \tag{16}$$

In order to specify μ and v independently we need a further condition and choose the variance $\langle\Delta\sigma^2\rangle$ of (15) which turns out to be

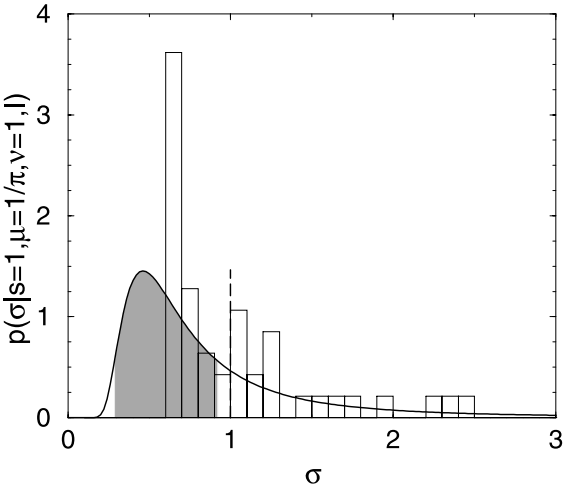

Fig. 1. Solid line: Probability distribution for the renormalized error σ. The mean $\langle\sigma\rangle = 1$ is given by the dashed line. The shaded area depicts the shortest 66% confidence interval. The histogram gives the distribution in the error correction for the data of Fig. 3.

$$\frac{\langle\Delta\sigma^2\rangle}{s^2} = \frac{\mu}{v-1} - 1. \tag{17}$$

We now choose to allow for large discrepancies between scattered data and specified errors and let $\langle\Delta\sigma^2\rangle/s^2$ tend to infinity. This fixes $v = 1$ and in consequence $\mu = 1/\pi$. We are now ready to eliminate the true errors σ from (11) by application of the sum rule (4) which allows us to reformulate the likelihood entirely in terms of measured data and model parameters. For a single term of the product (11) we obtain

$$\begin{aligned} P(\delta_i|x_i,s_i,c,\lambda,\mu,v,I) &= \int P(\delta_i,\sigma_i|x_i,s_i,c,\lambda,\mu,v,I)\,\mathrm{d}\sigma \\ &= \int P(\sigma_i|s_i,\mu,v,I) \\ &\quad\times P(\delta_i|\sigma_i,x_i,c,\lambda,I)\,\mathrm{d}\sigma, \end{aligned} \tag{18}$$

which may be readily integrated to yield with the above values of μ and v

$$P(\delta_i \mid x_i,s_i,c,\lambda,I) = \frac{1}{2\sqrt{\pi}}\frac{1}{s_i\sqrt{2\pi}}\Big/\left\{\frac{1}{\pi} + [\delta_i - c\varphi(x_i,\lambda)]^2/2s_i^2\right\}^{3/2}. \tag{19}$$

Similar expressions result in the case of (13). This is then the final form of our likelihood function which forms the basis for all our subsequent conclusions. Though our formulation is for reasons of simplicity only for data sets $\vec{\delta}$ and $\vec{\Delta}$ (with calibration parameter γ), the generalization to the case of multiple data sources $(\vec{\Delta}_k,\gamma_k)$ does not pose any difficulties.

4. Posterior estimates

In this section, we shall employ Bayes' theorem to infer parameters of the model functions, expectation values of the true errors of the measurements, and predictive values of the erosion, given flux, temperature and energy as well as the associated error. Let us begin with parameter estimation. From Bayes' theorem we obtain the posterior distribution of the parameters conditional on all experimental data

$$P\left(c,\lambda,\vec{\gamma}\middle|\vec{\delta},\vec{x},\vec{s},\vec{\Delta}_k,\vec{X}_k,\vec{S}_k,I\right) = \frac{P\left(c,\lambda,\vec{\gamma}\middle|I\right)P\left(\vec{\delta},\vec{\Delta}_k\middle|\vec{x},\vec{s},\vec{X}_k,\vec{S}_k,c,\lambda,\vec{\gamma},\mathbf{M},I\right)}{P\left(\vec{\delta},\vec{\Delta}_k\middle|\vec{x},\vec{s},\vec{X}_k,\vec{S}_k,\mathbf{M},I\right)}, \tag{20}$$

where $\vec{\Delta}_k$ stands for $\vec{\Delta}_1,\ldots,\vec{\Delta}_k$ ($\vec{X}_k,\vec{S}_k$ likewise). In order to carry on we now have to specify the prior $P(c,\lambda,\vec{\gamma}|I)$. Since all the variables are independent this factorizes to

$$P\left(c,\lambda,\vec{\gamma}\middle|I\right) = P(c|I)P(\lambda|I)\prod_{k=1}^{K}P(\gamma_k|I). \tag{21}$$

Our prior knowledge about c and λ is limited to the positivity constraint and we therefore use flat priors extending over the ranges $0\leqslant c\leqslant c_{\max}$ and $0\leqslant\lambda\leqslant\lambda_{\max}$. For the choice of γ_k we are not that ignorant, since a priori we can assume that each measurement is performed under best experimental conditions and should be assigned with an expectation value for the calibration factor of $\langle\gamma_k\rangle = 1$, where $0\leqslant\gamma_k<\infty$. The maximum entropy distribution under these assumptions is an exponential function

$$P(\gamma_k|I) = \mathrm{e}^{-\gamma_k}. \tag{22}$$

This completes the specification of the posterior distribution of the parameters given all the data and we can proceed to the evaluation of the expectation values and associated variances of $c,\lambda,\gamma_1,\ldots,\gamma_K$. Let ξ be any of $c,\lambda,\gamma_1,\ldots,\gamma_K$. We then have

$$\langle\xi^n\rangle = \int \mathrm{d}\lambda\int \mathrm{d}c\int \mathrm{d}^K\gamma\,\xi^n \times\left\{\frac{\mathrm{e}^{-\sum_{k=1}^{K}\gamma_k}P\left(\vec{\delta},\vec{\Delta}_k\middle|\vec{x},\vec{s},\vec{X}_k,\vec{S}_k,c,\lambda,\gamma,\mathbf{M},I\right)}{c_{\max}\lambda_{\max}P\left(\vec{\delta},\vec{\Delta}_k\middle|\vec{x},\vec{s},\vec{X}_k,\vec{S}_k,\mathbf{M},I\right)}\right\} \tag{23}$$

from which we calculate in particular $\langle\xi\rangle$ and, in order to determine the variance, $\langle\xi^2\rangle$. The bracketed term under the integral (23) is a normalized density in c, λ, $\vec{\gamma}$, which we use as the sampling density to evaluate the expectation values $\langle\xi^n\rangle$ by Markov Chain Monte Carlo (MCMC) techniques. Note that it is unimportant at this step whether we know the normalization (which we do not!) or not when applying the Metropolis Hastings algorithm for the MCMC.

Having obtained posterior values of the parameters and the associated errors we could proceed and compare the model function with the parameters fixed to their posterior expectation values $\hat{c}$, $\hat{\lambda}$, $\hat{\vec{\gamma}}$, to the experimental data. This is instructive but it is not the posterior inference on the erosion yield given flux, sample temperature, particle energy and kind. What we really want is the expectation value of the erosion yield $\langle Y(\vec{Z})\rangle$, where $\vec{Z}$ summarizes the experimental conditions, and of course the uncertainty associated with this estimate. The respective moments are calculated from the distribution $P(Y|\vec{Z},\vec{\delta},\vec{x},\vec{s},\vec{\Delta}_k,\vec{X}_k,\vec{S}_k,I)$. The latter is obtained using the sum rule as

$$\begin{aligned}P\left(Y\middle|\vec{Z},\vec{\delta},\vec{x},\vec{s},\vec{\Delta}_k,\vec{X}_k,\vec{S}_k,I\right) &= \int P\left(Y,c,\lambda,\vec{\gamma}\middle|\vec{Z},\vec{\delta},\vec{x},\vec{s},\vec{\Delta}_k,\vec{X}_k,\vec{S}_k,I\right)\mathrm{d}c\,\mathrm{d}\lambda\,\mathrm{d}^K\gamma\\ &= \int P\left(c,\lambda,\vec{\gamma}\middle|\vec{\delta},\vec{x},\vec{s},\vec{\Delta}_k,\vec{X}_k,\vec{S}_k,I\right)\\ &\quad\times P\left(Y\middle|\vec{Z},c,\lambda,I\right)\mathrm{d}c\,\mathrm{d}\lambda\,\mathrm{d}^K\gamma.\end{aligned} \tag{24}$$

The first factor under the integral (24) is already known from above (20)–(22) while the second is simply a delta function

$$P\left(Y\middle|\vec{Z},c,\lambda,I\right) = \delta(Y - c\varphi(x,T,E,\lambda)), \tag{25}$$

where we have explicitly resolved $\vec{Z}$ into flux x, sample temperature T and particle energy E. The δ-function form of (25) implies a complicated distribution (24). This complexity is unimportant since we are only interested in the expectation value $\langle Y\rangle$ of Y and the associated variance $\langle\Delta Y^2\rangle$. Performing the Y-integration to obtain the moments first, we obtain exactly formula (23) with $\xi\equiv c\varphi(x,T,E,\lambda)$. This is of considerable practical value since it means that $\langle Y\rangle$ and $\langle\Delta Y^2\rangle$ may be obtained in the same Monte Carlo run as the parameters c, λ, $\vec{\gamma}$. The full curves and the shaded bands in 3 and 4 display predicted erosion yields and associated confidence ranges for conditions given in the captions. It is important for applications to note that the correct Bayesian prediction of the erosion $\langle Y\rangle$ coincides with the model function evaluated at the posterior values of the parameters within error margin. For further discussion see Section 4 below.

Next we focus our attention on the measurement errors. In order to infer estimates of the true measurement error σ_i (or $\Sigma_{j,k}$) we need the distribution $P(\vec{\sigma},\vec{\Sigma}_k|\vec{D},I)$, where $\vec{D}$ summarizes $(\vec{\delta},\vec{x}),(\vec{\Delta}_1,\vec{X}_1),\ldots,$

$(\vec{\Delta}_K, \vec{X}_K)$, which by virtue of Bayes theorem may be written as

$$P\left(\vec{\sigma}, \vec{\Sigma}_k \middle| \vec{D}, I\right) = \frac{P\left(\vec{\sigma}, \vec{\Sigma}_k \middle| I\right)}{P\left(\vec{D} \middle| \mathrm{M}, I\right)} P\left(\vec{D} \middle| \vec{\sigma}, \vec{\Sigma}_k, \mathrm{M}, I\right). \tag{26}$$

We employ marginalization in order to recover the original likelihood

$$\begin{aligned} P\left(\vec{D} \middle| \vec{\sigma}, \vec{\Sigma}_k, \mathrm{M}, I\right) &= \int P\left(\vec{D}, c, \lambda, \vec{\gamma} \middle| \vec{\sigma}, \vec{\Sigma}_k, I\right) \mathrm{d}^K\gamma \, \mathrm{d}c \, \mathrm{d}\lambda / P(D|\mathrm{M}, I) \\ &= \int P\left(c, \lambda, \vec{\gamma} \middle| I\right) P\left(\vec{D} \middle| \vec{\sigma}, \vec{\Sigma}_k, c, \lambda, \vec{\gamma}, \mathrm{M}, I\right) \\ &\quad \times \mathrm{d}^K\gamma \, \mathrm{d}c \, \mathrm{d}\lambda / P(D|\mathrm{M}, I). \end{aligned} \tag{27}$$

We are now ready to calculate $\langle \sigma_i \rangle$ the posterior expectation value of the error of measurement δ_i by plugging together (27) and (26). We find that

$$\langle \sigma_i \rangle = \frac{2 s_i}{\sqrt{\pi}} \left\langle \left\{ \frac{1}{\pi} + \frac{(\delta_i - c\varphi(x_i, \lambda))^2}{2 s_i^2} \right\}^{1/2} \right\rangle, \tag{28}$$

where the average is again over the same density function as in (23).

In all the previous calculations $P(D|I) = P(\vec{\delta}, \vec{\Delta}_k \mid \vec{x}, \vec{s}, \vec{X}_k, \vec{S}_k, \mathrm{M}, I)$ entered formally, but we had no need to compute the numerical value of this function. Remember, however, that we started with two different models given by (12) and (14) and that comparison of these two models requires according to the odds ratio (6) the knowledge of just this quantity. A general procedure for likelihood functions of arbitrary structure has been given recently [22] and employed in [23]. Direct evaluation of the Bayes factor has also been described in the literature [24]. The likelihood in the present problem has a simple structure which suggests a simpler procedure. Simultaneously with all the other expectation values we calculate $\langle \xi_i \xi_k \rangle$ with $\vec{\xi} = (c, \lambda, \vec{\gamma})$ from (23) and thereby obtain the covariance matrix $\underline{\underline{C}}$ with elements

$$C_{ik} = \langle \xi_i \xi_k \rangle - \langle \xi_i \rangle \langle \xi_k \rangle, \tag{29}$$

which defines a Gaussian approximation to the sampling density in (23) via

$$G\left(\vec{\xi}, \underline{\underline{C}}\right) = G_0 \exp\left\{ -\frac{1}{2} \sum_{i,k} \frac{(\xi_i - \langle \xi_i \rangle)(\xi_k - \langle \xi_k \rangle)}{C_{ik}} \right\}. \tag{30}$$

G_0 is the value of the likelihood at $\langle \vec{\xi} \rangle$, the posterior estimate of the parameters. In this approximation the prior predictive value or global likelihood becomes

$$P\left(\vec{\delta}, \vec{\Delta}_i \middle| \vec{x}, \vec{s}, \vec{X}_i, \vec{S}_i, \mathrm{M}, I\right) = G_0 (2\pi)^{(K/2)+1} |\det \underline{\underline{C}}|^{1/2}. \tag{31}$$

There is a simple way to improve over this approximation [25]. From (23) taking $n = 0$ it follows that

$$\begin{aligned} P\left(\vec{\delta}, \vec{\Delta}_i \middle| \vec{x}, \vec{s}, \vec{X}_i, \vec{S}_i, \mathrm{M}, I\right) &= \int \mathrm{d}\lambda \int \mathrm{d}c \int \mathrm{d}\gamma^K \\ &\quad \times \left\{ \frac{\mathrm{e}^{-\sum_{k=1}^K \gamma_k} P\left(\vec{\delta}, \vec{\Delta}_k \middle| \vec{x}, \vec{s}, \vec{X}_k, \vec{S}_k, c, \lambda, \vec{\gamma}, \mathrm{M}, I\right)}{c_{\max} \lambda_{\max} G^*\left(\vec{\xi}, \underline{\underline{C}}\right)} \right\} \\ &\quad \times G^*\left(\vec{\xi}, \underline{\underline{C}}\right). \end{aligned} \tag{32}$$

The bracketed term on the right-hand side of (32) is now expected to be a smoothly varying function of c, λ, $\vec{\gamma}$ and the integral may be performed by sampling from the normalized Gaussian $G^*(\vec{\xi}, \underline{\underline{C}})$. In the present work we have employed the Gibbs sampler [24] since it is simple and easy to derive full conditional distribution from a multivariate Gaussian. With the obtained values for the global likelihoods (32) we can finally evaluate the odds ratio (6). The results are presented in Section 6.

5. Empirical model function

We now come to the assignment of the yield function $Y_{\mathrm{tot}} = c\phi_0(\Phi, \varepsilon)$. For a long time [26–28] the decrease in yield with increasing flux was empirically described by an inverse function with power ε, i.e.,

$$Y_{\mathrm{tot}}(c, \varepsilon) = c \left(\frac{\Phi}{\Phi_0} \right)^{-\varepsilon}, \tag{33}$$

where Φ_0 is a normalization constant in order to have dimensionless units. This implies that the erosion would ever increase for vanishing flux. However, the erosion yield does saturate for smaller fluxes and finally becomes constant. Therefore, an empirical model function should rather read

$$Y_{\mathrm{tot}}(c, \varepsilon) = c \frac{1}{1 + \left(\frac{\Phi}{\Phi_0} \right)^{\varepsilon}}. \tag{34}$$

In the next section we will use a sophisticated model obtained from analytical considerations with the same flux dependence, but without an additional exponent ϵ. Our interest in this section is therefore to clarify, if a fit to the data with Eq. (34) does indicate the necessity of this further parameter.

In order to achieve this three different data sets are examined. The expectation values of the model parameters (ε, Φ_0) are calculated within Bayesian probability theory, while the value of c is determined from the low

flux data provided by the high current source (HCS [16]). No recalibrations of the data sets are taken into account. The result for the first data set, which comprises deuterium data from HCS, PSI-1 [12] and PISCES [13], are shown in Fig. 2. The exponent computes to $\langle\varepsilon\rangle = 0.79 \pm 0.12$. The pronounced flux dependence of the erosion yield results mainly from the HCS and PSI-1 data. Due to the large error bar of the PISCES data it is not possible to draw any conclusions on a functional dependence, although they still contribute in defining a range where the model function has to pass through. As well as for the second data set (deuterium data from HCS and the divertor region of ASDEX Upgrade, not shown) $\langle\varepsilon\rangle = 0.90 \pm 0.21$ and the third (hydrogen data from HCS and ASDEX Upgrade, not shown) $\langle\varepsilon\rangle = 1.5 \pm 0.3$, the exponent results in values close to one. So from this results we argue, that there is fortunately no need for an additional exponent assigned to the yield dependence on the flux, since any value different from $\varepsilon = 1$ would cause severe interpretation difficulties.

Though more data are available from the HCS in the first data set, we could only focus on those which have the same energy of the incident ions as the PSI-1 and PISCES data, i.e., $E_0 = 30$ eV. Even worse, the same restriction reduces, e.g., in the second data set the number of the available ASDEX Upgrade data to one fourth of the total set. A far better way is therefore to incorporate the incident ion energy in a model which considers the physical processes of the erosion and thus

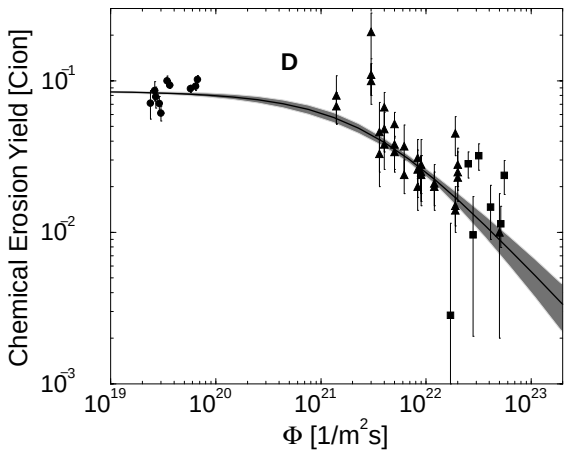

Fig. 2. The chemical erosion yield obtained by a fit to the empirical model of Eq. (34). Surface temperatures are around $T_{\max}$. The circles are data from the high current source [16], triangles depict the PSI-1 [12] data and squares are due to the data from PISCES [13]. Error bars show the assigned experimental error. The solid line represents a fit with $\varepsilon = 0.79$ ($\Phi_0 = 3.2 \times 10^{21}$, $c = 0.0856$). The gray shaded area is the confidence range.

allows to include data taken under slightly different conditions (e.g., E_0 between 20 and 50 eV, temperatures between 300 and 350 K) in the determination of the posterior distribution. This is done in the next section.

6. Results and discussions of the full model

The described mathematical treatment is applied to the evaluations of the flux dependence of chemical erosion of graphite under hydrogen irradiation. The set δ was chosen to be the data obtained from ion beam irradiation of carbon samples and weight loss measurement of the target [16]. These data are to be considered as correctly calibrated, since the measurement method is the most simple one and produces reliable total yields. At large ion fluxes three sets of data were chosen, either from plasma simulators [12,13] or from measurement in the divertor of the fusion experiment ASDEX Upgrade [14] where the eroded molecule flux was determined spectroscopically from the CH band intensity. In view of the uncertainties in the evaluation of the eroded CH_4 or C_2H_n flux from the CH band intensity, a calibration parameter γ_k was introduced as fitting parameter (model M_1). A second approach assumed a calibration error of the spectrometer (model M_2) and uses the same calibration parameter γ_k also to the determination of the incident hydrogen flux obtained spectroscopically from the H_α line. Both choices for the calibration parameter γ_k are finally examined in the model comparison.

As hypothesis H in the Bayesian treatment the formulae for the chemical erosion yield Y_{tot} of [7] have been taken which include the $1/\Phi$ flux dependence at high ion fluxes as obtained from the free fit to the data. The range of energies considered in the present treatment allows to neglect physical sputtering otherwise included in the weight loss data. The formulae including the fitting parameters c and $\lambda = 1/\Phi_0$ read schematically

$$Y_{\text{tot}}(c,\lambda) = cY_{\text{chem}}(E,T,\Phi)\frac{1}{1+\lambda\Phi}. \tag{35}$$

The flux dependence of Y_{chem} is very weak in comparison with the last term. This last term gives the concentration of hydrocarbon radicals at the surface and results from the assumption that the hydrogenation of surface atoms to radicals, which can subsequently be eroded under ion irradiation, is limited by a time constant τ. The parameter λ, which characterizes the flux where the week flux dependence of Y_{tot} given by Y_{chem} turns into a $1/\Phi$ flux dependence, is dependent on the rate at which radicals are removed from the surface and the hydration time τ.

Figs. 3 and 4 show the resulting fit of the model function to the available data. Fig. 3 shows data obtained at temperatures close to the maximum of chemical erosion, $T_{\max}$, in the energy range 20–50 eV.

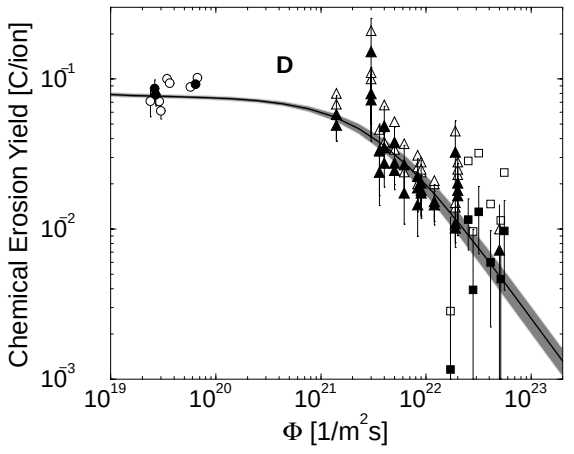

Fig. 3. Dependence of the chemical erosion yield on ion flux Φ for surface temperatures around T_{max}. The circles are data from the high current source, where the full symbols represent the subset corresponding to $E_0 = 30$ eV. The experimental outcome of PSI-1 is given by open triangles. The data were corrected by a factor of $\gamma_1 = 0.83$ to result in the filled triangles. The same representation was chosen for the PISCES data (open and filled squares), where the factor is $\gamma_2 = 0.41$. Error bars show the assigned experimental error. The Bayesian result, shown by the solid line is the fit of the yield for $E_0 = 30$ eV and a temperature of 600 K. The gray shaded area is the confidence range.

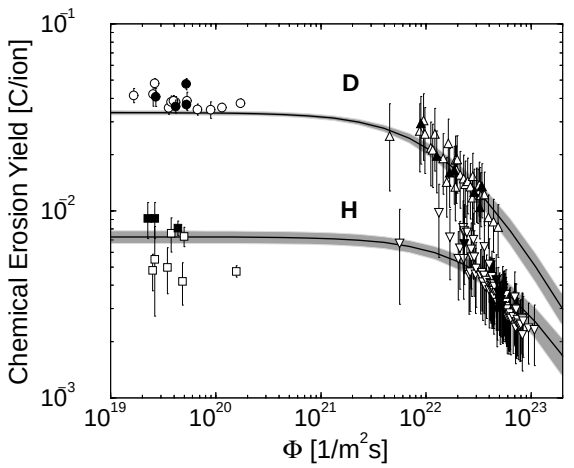

Fig. 4. The same as Fig. 3, but for T around room temperature. The upper (lower) curve represents erosion with deuterium (hydrogen) [14]. Again the circles (squares) are data at $T = 300$ K from the high current source, where the full symbols represent the $E_0 = 30$ eV subset. The ASDEX data at $T = 350$ K are depicted by triangles showing up (down). Here already the renormalized data are shown, where the full symbols represent 25 eV $< E_0 <$ 35 eV. The correction factor is $\gamma = 1.13$ ($\gamma = 0.97$). Both fits (solid line) were obtained for $E_0 = 30$ eV and a temperature of 350 K.

Spectroscopic data at high ion flux was always at $E_0 = 30$ eV and the evaluation has been made for $E_0 = 30$ eV (solid data points for ion beam data) while the evaluation of points deviating from 30 eV (open data points for ion beam data) took the energy dependence of Y_{chem} into account [7]. Open symbols for the spectroscopic data denote the published values Δ_k, full symbols after multiplying with the calibration parameter γ_k. For a judgement of the quality of the fit (solid curve in Figs. 3 and 4) only the full symbols should be considered. Though the full Bayesian approach giving the expectation value for Y_{tot} is the correct solution, it is important to note that the model function evaluated with the posterior values of the parameters provides us with a hardly distinguishable result. This is due to the large number of data and simplifies the situation a lot, since one can calculate the total yield from the model function by just inserting $\langle c \rangle$ and $\langle \lambda \rangle$. With these two values and the use of model function (12) it is now possible to compute any Y_{tot} in the supported flux range (10^{19}–10^{24}/m^2 s), a fact which is of particular interest for the modeling of carbon erosion in ITER. Table 1 gives the results for these parameters and the calibration factors γ_k as well as the odds ratio, distinguishing between the two models in applying the calibration parameter γ_k as outlined in section 6. This odds ratio does not deviate largely from unity, and did not give any reason of deferring from the statement of the experimentalists, that the correction factor should be applied to the eroded atom flux only, rather than to the incident hydrogen flux and eroded atom flux. Therefore, the results depicted in Figs. 3 and 4 refer to the first model (12), where the calibration factor is only applied to the yield data, $\gamma_k \Delta_k$. Since the calibration factor γ_k provides the model with a further degree of freedom, the linear factor c is mainly determined by the ion beam data $\vec{\delta}$ which were taken to be correctly scaled. The calibration factors are 0.83 for the data from [12] and 0.41 for [13]. The value of λ resulting from the fit is 22.2×10^{-23} m^2 s. Fig. 4 shows the same treatment, but for data around room temperature ($T \approx 350$ K). The upper curve represents erosion with deuterium, the lower with hydrogen ions. Since the calibration parameter γ is close to unity (1.13 for D and 0.97 for H) here only the renormalized data are shown for the spectroscopic data at high flux. Full symbols represent data at $25 < E_0 < 35$ eV while for the treatment of data at other energies (open data points) the energy dependence of Y_{chem} was taken into account. Both fits (solid line) were obtained for $E_0 = 30$ eV and $T = 350$ K. The values of λ are 5.4×10^{-23} m^2 s for D and 1.65×10^{-23} m^2 s for H.

An interesting point is the dependence of λ on the erosion yield at low ion flux Y_{low}. Here the number of surface radicals is close to the areal density A of target atoms within the ion range and λ can be written as

Table 1
Model parameter and calibration factor γ for the data sets of Figs. 3 and 4. The true error $\langle\sigma\rangle$ is in good agreement to the estimated error s[a]

	$\langle c\rangle$	$\langle\lambda\rangle$ (10^{-23} m^2 s)	$\langle\gamma\rangle$	$\overline{\sigma}/s$	Odds ratio
HCS, PSI-1, PISCES	3.05	22.2	0.83, 0.41	0.97	1.5
HCS, ASDEX(D)	1.93	5.42	1.13	1.0	1.1
HCS, ASDEX(H)	1.18	1.65	0.97	0.92	0.72

[a] The odds ratio $P(M_1|D,I)/P(M_2|D,I)$ shows no strong discrimination between models M_1 and M_2.

Table 2
Hydration time τ as calculated from λ and Y_{low} at low ion flux according to Eq. (36)[a]

	λ (10^{-23} m^2 s)	Y_{low} (C/ion)	τ (10^{-3} s)
HCS, PSI-1, PISCES	22.2 ± 5.4	0.079 ± 0.002	0.28 ± 0.07
HCS, ASDEX(D)	5.4 ± 1.2	0.0335 ± 0.0008	0.16 ± 0.04
HCS, ASDEX(H)	1.65 ± 0.39	0.0073 ± 0.0006	0.23 ± 0.06

[a] The mean gives $\bar{\tau} = 0.22 \pm 0.03$.

$$\lambda = \frac{Y_{low}}{A}\tau. \tag{36}$$

The values of λ and Y_{low} are given in Table 2 together with the resulting hydration times τ. An areal density of target atoms within the ion range of about $10^{17}/\text{cm}^2$ was assumed. The hydration times for different surface temperatures and different hydrogen isotopes agree well within the error margins and therefore allow to compute a mean of $\bar{\tau} = 0.22 \pm 0.03$ *m*s. This appears rather long, but is in good agreement with the estimate of 1 m s obtained from [11].

Finally, for all sets of data $(\vec{\delta}, \vec{\Delta}_k)$ estimations of the associated $(\vec{s}, \vec{S}_k)$ were provided by the experimentalists. After the evaluation of the posterior errors employing (28) it is possible to compare them to the initial probability assignment. The posterior error distribution is shown as the histogram in Fig. 1. Of course, by design the average value of the histogram is the same as for the continuous curve namely 1. However, also the initial choice of an infinite variance (17) seems to be corroborated by the data.

7. Conclusion

The evaluation of the rather limited data at high ion flux reveals indeed, that a strong flux dependence sets in around ion flux of the order of 10^{21}–$10^{22}/\text{m}^2$ s. The confidence level of the fit of the model to the data is within 30% for recalibration parameters only marginally deviating from unity. Only one data set [13] required a renormalization by a factor of 0.41, but data points at the highest flux may alternatively indicate an additional, flux independent erosion mechanism which was not considered here. The discussion of the dependence of the values of λ on the erosion yield at low ion fluxes resulted in a hydration constant of 0.22 m s independent on surface temperature or hydrogen isotope.

Acknowledgements

We thankfully acknowledge valuable discussions with the authors of high flux data for chemical erosion, especially on the origin of the published experimental error bars. In the case of ASDEX Upgrade an estimate of experimental errors was provided by A. Kallenbach.

References

[1] A. Kallenbach et al., Nucl. Fus. 34 (1994) 1557.
[2] H.Y. Guo et al., J. Nucl. Mater. 241–243 (1997) 385.
[3] W. Poschenrieder et al., J. Nucl. Mater. 220–222 (1995) 36.
[4] H. Kubo et al., J. Nucl. Mater. 196–198 (1992) 71.
[5] H. Verbeek, J. Stober, D.P. Coster, W. Eckstein, R. Schneider, Nucl. Fus. 38 (1998) 1789.
[6] G. Federici et al., J. Nucl. Mater. 266–269 (1999) 14.
[7] J. Roth, J. Nucl. Mater. 266–269 (1999) 51.
[8] A. Horn et al., Chem. Phys. Lett. 231 (1994) 193.
[9] J. Roth, C. García-Rosales, Nucl. Fus. 36 (1996) 1647 with corrigendum Nucl. Fus. 37 (1997) 897.
[10] J.W. Davis, A.A. Haasz, J. Nucl. Mater. 241–243 (1997) 37.
[11] J.W.A.M. Gielen, M.C.M. van de Sanden, D.C. Schramm, Appl. Phys. Lett. 69 (1996) 152.
[12] H. Grote et al., J. Nucl. Mater. 266–269 (1999) 1059.
[13] D.G. Whyte, G.R. Tynan, R.P. Doerner, J.N. Brooks, Nucl. Fus. 41 (2001) 47.
[14] A. Kallenbach, A. Thoma, A. Bard, K.H. Behringer, K. Schmidtmann, M. Weinlich, ASDEX Upgrade team, Nucl. Fus. 38 (1998) 1097.
[15] R.D. Monk et al., Phys. Scripta T81 (1999) 54.
[16] M. Balden, J. Roth, J. Nucl. Mater. 280 (2000) 39.
[17] P.C. Gregory, T.J. Loredo, Astrophys. J. 398 (1992) 146.

[18] G. Larry Bretthorst, Bayesian Spectrum Analysis and Parameter Estimation, Springer, Berlin, 1988.
[19] A. O'Hagan (Ed.), Kendall's Advanced Theory of Statistics, Bayesian Inference, Wiley, New York, 1994.
[20] D.S. Sivia, Data Analysis, Oxford University, Oxford, 1996.
[21] V. Dose, W. von der Linden, Outlier tolerant parameter estimation, in: Maximum Entropy and Bayesian Methods, Kluwer Academic, Dordrecht, 1999.
[22] W. von der Linden, R. Preuss, V. Dose, The prior predictive value: a paradigm of nasty multi-dimensional integrals, in: Maximum Entropy and Bayesian Methods, Kluwer Academic, Dordrecht, 1999.
[23] R. Preuss, V. Dose, W. von der Linden, Nucl. Fus. 39 (1999) 849.
[24] R.Neal, Probabilistic inference using Markov Chain Monte Carlo methods, Department of Computer Science, University Toronto, 1993.
[25] Ruanaidh, W. Fitzgerald, Numerical Bayesian Methods Applied to Signal Processing, Springer, New York, 1996.
[26] S.K. Erents et al., J. Nucl. Mater. 63 (1976) 399.
[27] J. Roth, E. Vietzke, A.A. Haasz et al., in: Atomic and Plasma-Material Interaction Data for Fusion (Suppl. to Nucl. Fus.), vol. 1, IAEA, Vienna, 1991, p. 63.
[28] J.W. Davis, A.A. Haasz, J. Nucl. Mater. 241–243 (1997) 37.