Smoothing Methods for Continuous Permeation Data Measured Discretely Designated for Quick Evaluation of Barrier Materials

K. Hasilová[*1], P. Otřísal[2] and P. Stodola[3]

[1]*Department of Quantitative Methods, University of Defence, Czech Republic*
[2]*Faculty of Physical Culture, Palacký University Olomouc, Czech Republic*
[3]*Department of Intelligence Support, University of Defence, Czech Republic*

Abstract:

Information concerning barrier properties of materials used for production of personal protective equipment fundamentally not only affects their useful properties, but also supports the end users' decisions. Data obtained from measurements by standardized methods have to be processed by appropriate statistical methods. The article deals with numerical and statistical methods of reconstruction of a continuous curve out of discretely measured data. These mathematical models are proposed as an extension to the manual measurement of a conductometric method. Behaviour of these models is demonstrated on two chemical warfare agent simulants, however, the proposed methodology is universal and can be applied also to chemicals which have no conductivity. The parametric and nonparametric models are studied in the curve reconstruction, as well as in the calculation of subsequent characteristics, for example, a lag-time. The nonparametric model shows the best results which are in accordance with an expert's estimate.

Keywords:

1 Introduction

Permeable and insulating barrier materials are designed to protect the user's body surface against the effects of toxic substances. Their basic purpose is to separate the outdoor area with the presence of a mass concentration of chemical warfare agents (CWA) or toxic industrial chemicals (TIC) from the area where the toxic substance cannot appear. In

---

* *Corresponding author: Department of Quantitative Methods, University of Defence, Kounicova 156/65, CZ-662 10 Brno, Czech Republic. E-mail: kamila.hasilova@unob.cz. ORCID 0000-0003-1540-3489.*

this context, it should be understood that a very small amount of CWA can cause irreversible loss or even the death of the protective device user. The long-term protection against the permeation of toxic substances is thus determined by the concentration of the toxic substance and the type of used barrier material. Due to the different chemical and physical-chemical characteristics of CWA and TIC, it is currently not possible to find such a breathable or insulating barrier material that meets high chemical resistance requirements while meeting other requirements associated with the long-term activity of military forces in the contaminated area [1-5]. It will always be a compromise between the provided level of protection and the ability to perform an operational task [6-8].

Methods that have been used in the recent time for determination the chemical resistance of barrier materials against permeation of toxic substances have been largely based on the principle of colorimetric reaction and subsequent colour response of the detection layer. Although these methods were accurate enough and provided very good and realistic data concerning the breakthrough time of a particular barrier material, they did not provide the data required by European standards. In the case of colorimetric methods or methods based on a chemical principle, the breakthrough time is expressed as the time from the beginning of contamination of the barrier material with a particular CWA or TIC to the point where the first detectable, thus visible change was registered. This usually corresponds to the so-called threshold dose. In the case of CWA, the data were available because they were surveyed in a targeted and long-term manner. However, in the context of a change in the security environment within the North Atlantic Treaty Organization and a different specification of security threats, it was necessary to develop methods and methodologies that would allow the assessment of barrier materials against the permeation of TIC where information related to threshold doses is completely absent [9-11].

Currently, there is a growing demand for rapid and efficient information sharing in all types of military operations. Commanders' and staffs' information support is directed not only to provide information and data concerning classical combat operations of the units, but also to provide information pointing to the level of provided antigas protection. If there is a possibility to quickly assess the ability of barrier materials to resist contamination caused by CWA or TIC, the commander can make an immediate decision to adjust the regime of work in the contaminated area, and to prepare preventive measures to ensure rest and rotation of soldiers and so on.

Aim of the Contribution

To evaluate the permeation of toxic chemicals through the barrier material, the values shown in the cumulative amount versus time graph or the CWA or TIC concentration versus time graph are essential. The part of the permeation curve from which the steady-state permeation rate can be read is critical for determining the chemical resistance value using lag-time. The currently used permeation calculator [12] makes it possible to read the value of the breakthrough time; however, only by designating manually a specific place on the permeation curve. The lag-time value is then read from the linear part of the graph of the permeation mass depending on time. This reading is also done manually by clicking on the graph to enter a line that is tangent to the linear part of the graph and its intersection with the time axis determines the lag-time value.

Since the measurement usually takes place at set time intervals, the obtained values of measurement of the above mentioned characteristics are in the form of discrete recordings of continuous quantities. The first step is to find a continuous function that appropriately

represents the empirical data. Assuming such a function exists, the function can be used for modelling and it can be further developed, which is better than working with a set of discrete values. In particular, it is possible to calculate the value of the breakthrough time exactly and to construct a tangent line of the function analytically as well. In such cases, the application of smoothing methods, either parametric or nonparametric, seems to be very suitable.

The main objective of this article is to provide a basis for an automation of the measuring procedure so that it is not necessary to determine values from the permeation curve manually. Sub-objectives are then to propose an optimal way of estimating a continuous function, which is hidden in the measured discrete values [13-16] and to provide a mathematical-statistical background for calculations of the residence time and other permeation characteristics acquired from the permeation profile.

2 Data and Methods

A number of methods are currently used to test the chemical resistance of barrier materials. These methods work on different analytical principles. In this article, we focus on methods that are based on conductivity.

One of the variables that is useful as a starting point for the evaluation of the permeation of toxic substances through barrier materials is conductivity. The change of conductivity corresponds to the concentration of dissociated ions of used test chemical in a well-defined volume of redistilled water. After calibration of the measuring electrode response to a specific and very precise amount (volume) of the used test chemical, it is possible to obtain information about the change in conductivity of the specific exact volume of the redistilled water. Within the practical implementation of the permeation measurement, the used test chemical, which is diffused through the tested barrier material, is absorbed into the redistilled water. After dissociation, the conductivity is determined and converted to concentration by the permeation calculator. When measuring the permeation of toxic substances under static conditions, the concentration of the test chemical in the redistilled water gradually increases, and so does the concentration of measurable dissociated ions [17].

KONDUKTOTEST is a method for the quick determination of the resistance of porous (filtration) and non-porous (insulation) barrier materials to the static permeation of sulphur mustard and other volatile toxic compounds with acidobasic properties which are soluble in water on ion dilution. The KONDUKTOTEST device, which was developed at the Technical Research Institute of Protection in Brno (Czech Republic), is a semi-automatic device enabling continuous monitoring of the permeation of volatile toxic substances through the materials of insulating and filtering protective equipment using a conductivity sensor consisting of carbon electrodes [18].

The use of the KONDUKTOTEST method fully respects the requirements based on the standard classical analytical method of conductivity. The general principle of conductometry is applied, however, the test chemical dissociates into ions only after it has penetrated the barrier material being tested. This leads to a more versatile approach in the application of a wide range of methods useful for barrier material testing.

For the purpose of evaluation of the resistance of barrier materials against the permeation of toxic substances, the start of the permeation curve (see Fig. 1), which corresponds

to the state where no toxic substance has permeated through the barrier material, is essential. This corresponds to the time when the toxic substance was not dissociated in the redistilled water, thus no response of the measuring system was recorded. The linear part of the graph corresponds to the state where the toxic substance already permeates intensively through the tested material and is used as a tool for reading the lag-time value. For the purpose of evaluation of the resistance of barrier materials against permeation of toxic substances, this is the time when the amount of the dissociated test chemical is detected that would cause death or disqualification of a user of the antigas protective equipment. The last part of the graph, thus reaching steady state permeation rate, expresses the limit of the detection system to react to the detected amount of the permeated toxic substance. It has no longer any practical significance for the purposes of assessing the resistance of barrier materials under military conditions. Achieving this state indicates a situation where the permeation measurement can be terminated.
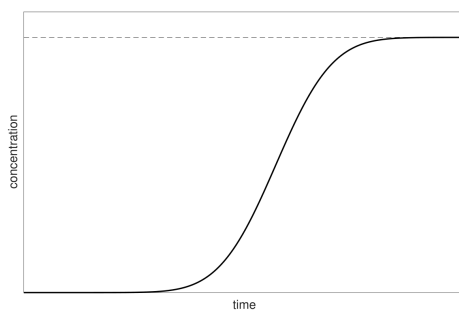

Fig. 1 Typical course of the concentration curve in time

2.1 Data

Chemical warfare agents are chemical compounds and mixtures which within their use in combat operations can kill, seriously injure or incapacitate persons, contaminate the environment, people, weapons, objects and other material. Due to the very high toxicity and hazardousness of this type of chemicals and the necessity to respect very strict safety measures related to health protection, it is advisable to use substitute test substances whose behaviour is as close as possible to the behaviour of CWA. In this study, we have used two substitute substances, namely cyclohexylamine and pyridine.

Cyclohexylamine is an organic compound belonging to the class of aliphatic amines. It is a colourless liquid. Samples are often coloured due to chemical impurities. However, colour (chemical purity) has no effect on permeation measurements. Cyclohexylamine is characterised by its fishy smell. However, for the purpose of conductivity measurements it is very important that it is infinitely miscible with water, which means that it can form aqueous solutions with a very high conductivity response. In practice, this means that in an aqueous environment, molecules dissociate very rapidly into ions, which can then be detected by a conductivity measuring device. Such behaviour corresponds to the characteristics of those CWA and TIC that have the same or similar chemical properties, thus

those that dissociate unrestrictedly in water. For this reason, cyclohexylamine is considered a standard test chemical for the evaluation of barrier properties in relation to those toxic substances that dissociate unrestrictedly in the aquatic environment.

Pyridine is an organic aromatic heterocyclic compound of carbon, hydrogen and nitrogen. Under normal conditions, it is a characteristically odourless, colourless, flammable liquid miscible with water and ethanol. In this case, however, unlimited miscibility is not associated with an unlimited ability to dissociate into ions detectable by the measuring conductometric device. The use of this type of chemical is based on the fact that some CWA and TIC dissociate in aqueous environments only to a limited extent, thus, in quantities and concentrations approximately equivalent to the dissociation capacity of pyridine. For this reason, pyridine is considered a standard test chemical substance for the evaluation of barrier properties in relation to those toxic substances that dissociate only to a limited extent in the aquatic environment.

2.2 Models of the Measured Data

Regression models provide a suitable approach to modelling the data, i.e., modelling the dependence of concentration on time. A standard regression model is defined as

$$c_i = F(t_i, \boldsymbol{a}) + \varepsilon_i, \quad i = 1, \dots, n$$

where c_i is the concentration of the chemical substance at fixed time points t_i, F is the regression function, $\boldsymbol{a} = (a_1, \dots, a_k)$ denotes a vector of unknown parameters, and ε_i are independent and identically distributed random variables [19].

Parameters $\boldsymbol{a}$ of the model are selected using the least square approach, i.e., minimising the square of the vertical distance between the measured concentration values and the estimated values

$$\boldsymbol{a} = \arg\min_{\boldsymbol{a} \in \mathbb{R}^k} \left[c_i - F(t_i, \boldsymbol{a}) \right]^2$$

Piecewise Linear Model

Since the focus of the research is on the initial and middle part of the curve, the simplest piecewise linear estimate of the following form could be used:

$$c(t) = a_1 + a_2 \cdot (t - z)_+ \tag{1}$$

where a_1, a_2 denote the parameters of the function and z is the breakpoint of the function. Function $(\cdot)_+$ is defined as

$$(A)_+ = \begin{cases} A & \text{if } A > 0 \\ 0 & \text{if } A \leq 0 \end{cases}$$

Parametric Model

Two models can be differed here, linear and nonlinear, depending on the linearity in the parameters. The linear model has one basic configuration; the nonlinear model is more flexible and opens door to a great number of possible forms. From nonlinear model candidates, we selected the logistic curve model, which is in accordance with the typical shape of the concentration curve. The model takes the form

$$c(t) = \frac{a_1}{1 + \exp[-a_2 \cdot (t - a_3)]}$$

where a_1 denotes the curve maximum value, a_2 is known as the logistic growth rate and a_3 denotes the curve midpoint [20].

Nonparametric Model

A nonparametric model can be written in the same simple form as the parametric model, i.e., $c(t) = F(t) + \varepsilon$. The shape of the function F draws a distinction between the parametric and nonparametric models. In the parametric approach, it is assumed that F has a prespecified functional form. On the other hand, nonparametric estimates attempt to reconstruct the unknown function from the data themselves using as few assumptions as possible.

Such a model can be fitted, among others, by using a kernel regression model with local linear polynomials. It involves solving the least square problem

$$\sum_{i=1}^{n} \left[c_i - b_0 - b_1(t - t_i) \right]^2 \cdot W_i(t, h) \longrightarrow \min \tag{2}$$

with c_i standing for the measured concentration values at times t_i, letters b_0, b_1 stand for the unknown parameters and $W_i(t, h)$ is a weight function that depends on parameter h [21].

The kernel regression model uses a function called kernel as the weight function. The kernel is usually considered to be a symmetric nonnegative function which integrates to one. Another parameter used in the kernel estimate is h, which is called a bandwidth; it influences the smoothness of the resulting estimate [22-24].

Specifically, we have opted for the weight function based on the Gaussian density

$$W_i(t, h) = \frac{1}{h} K\left(\frac{t - t_i}{h}\right), \quad K(x) = (2\pi)^{-1/2} \cdot \exp(-x^2/2)$$

and smoothing parameter h, which minimizes a quantity based on Kullback-Leibler cross-validation criterion [25, 26].

Then the estimate of the regression function at point t takes the form:

$$\widehat{F}(t) = \frac{1}{nh} \sum_{i=1}^{n} \frac{\left[s_2 - s_1 \cdot (t - t_i) \right] \cdot K\left(\frac{t - t_i}{h}\right) \cdot c_i}{s_2 \cdot s_0 - s_1^2}$$

where

$$s_j = \frac{1}{nh} \sum_{i=1}^{n} (t - t_i)^j \cdot K\left(\frac{t - t_i}{h}\right) \quad j = 0, 1, 2$$

Assessing the Quality of the Models

For assessing the goodness-of-fit of the proposed models, a criterion which allows to compare parametric and nonparametric models together has to be applied. Therefore, an alternatively defined coefficient of determination is used [27]. It allows comparing parametric and nonparametric estimates of the linear and nonlinear regression functions. Denoting $\widehat{c}_i$ the fitted value for the outcome c_i and $\overline{c}$ the mean of the outcomes, the coefficient of determination is defined as

$$R^2 = \frac{\left[\sum_{i=1}^{n} (c_i - \overline{c})(\widehat{c}_i - \overline{c}) \right]^2}{\left[\sum_{i=1}^{n} (c_i - \overline{c})^2 \right] \cdot \left[\sum_{i=1}^{n} (\widehat{c}_i - \overline{c})^2 \right]}$$

The value of R^2 lies in the interval from zero to one with the value 1 indicating the perfect fit.

2.3 Lag-time

The lag-time value is estimated as the intersection of the tangent line to the linear part of the concentration curve and the time axis. Regarding the general shape of the concentration curve, the linear part is in the neighbourhood of the inflection point, i.e., the point where the first derivative reaches its maximum value. (The minimum value of the first derivative, which also could lead to the inflection point, is not considered because of the overall behaviour of the data.) Therefore, we need to calculate or estimate the first derivative, find its maximum point, construct the tangent line, and find its intersection with the time axis.

The calculation of the first derivative for the piecewise and parametric model is straightforward since we know the formula of the respective functions. Namely, the first derivatives of the piecewise linear model parametric logistic model are

$$
\begin{aligned}
\text{linear} \quad & c'(t) = \begin{cases} 0 & \text{if } t < z \\ a_2 & \text{if } t > z \end{cases} \\
\text{logistic} \quad & c'(t) = \frac{a_1 \cdot a_2 \cdot \exp[-a_2 \cdot (t - a_3)]}{\{1 + \exp[-a_2 \cdot (t - a_3)]\}^2}
\end{aligned}
\tag{3}
$$

However, there is neither a close form of the kernel estimate of the concentration curve, nor its derivative. In general, there are two ways to estimate the derivatives of the kernel estimate:

  (i)   differentiate the estimate of the function,
 (ii)   estimate directly the derivative of the function.

The former way can be divided into two options. Namely, we can calculate the central differences of the estimated function,

$$
\widehat{F}'(t) = \frac{F(t + \Delta t) - F(t - \Delta t)}{2\Delta t}
$$

or we can use the derivative of the kernel estimate [28]. If the kernel estimate is written in its general form as $\widehat{F}(t) = n^{-1} \sum_{i=1}^{n} W_i(t,h) \cdot c_i$, then the derivative of the kernel estimate takes the form $\widehat{F}'(t) = n^{-1} \sum_{i=1}^{n} W_i'(t,h) \cdot c_i$, where the derivative is taken with respect to t.

The latter way of estimating the derivatives consists of two approaches. The first one is based on the generalised version of Eq. (2)

$$
\sum_{i=1}^{n} \left[c_i - \sum_{r=0}^{R} b_r \cdot (t - t_i)^r \right]^2 \cdot W_i(t,h)
\tag{4}
$$

If the values $b_0, b_1, \ldots, b_R$ minimize this equation, then the estimated r^{th} derivative of $F(t)$ is given by $\widehat{F^{(r)}}(t) = r!b_r$ for $r \leq R$ [29, 30]. The other approach is based on calculating the regression function not on the data themselves, but on their symmetric differences

$$
c_{ij}^{(1)} = \frac{c_{i+j} - c_{i-j}}{t_{i+j} - t_{i-j}}, \quad 1 \leq j \leq m
$$

where m is the positive integer [31-33].

3   Results and Discussion

In this section, a description of modelling the discretely measured continuous data and subsequent calculation of the lag-time is provided.

3.1 Concentration Curve Estimates

Estimates of the piecewise linear and logistic models and their parameters were found by minimization of the squared (vertical) distance between the model and the data using the Matlab function *fminsearch*, which is based on the Nelder–Mead algorithm for finding minima of a function of several variables. The bandwidth of the nonparametric model was calculated using the R function *npreg* (from the library *np*).

The proposed methods of modelling the discretely measured data are presented on two data sets—cyclohexylamine and pyridine—here denoted as data A and data B, respectively. In Figs 2 and 3, all the proposed estimates are displayed: the piecewise linear model in blue, the parametric model based on the logistic curve in orange and the nonparametric model represented by the kernel regression in green.
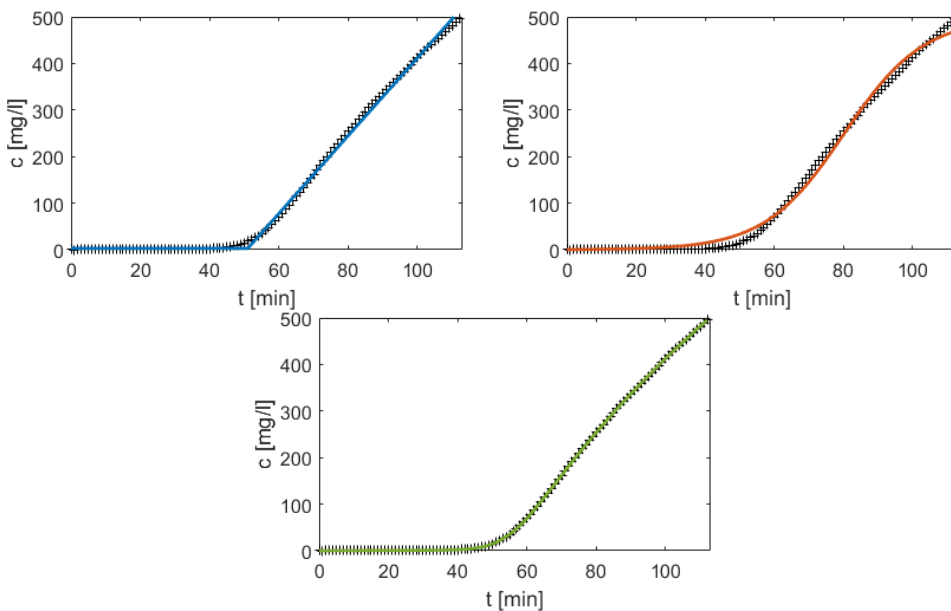

Fig. 2 All proposed estimates of the data set A data (black +), piecewise linear (blue), logistic (orange), kernel regression (green)

It can be easily seen that the nonparametric model gives estimates which are closer to the data than the parametric—the logistic and piecewise linear—models. As it is obvious from the graphs in Fig. 2, for data A, all of the proposed estimates give reasonable curves describing the behaviour of the concentration over time. On the other hand, only the nonparametric model can be assessed as appropriate for the data set B (Fig. 3).
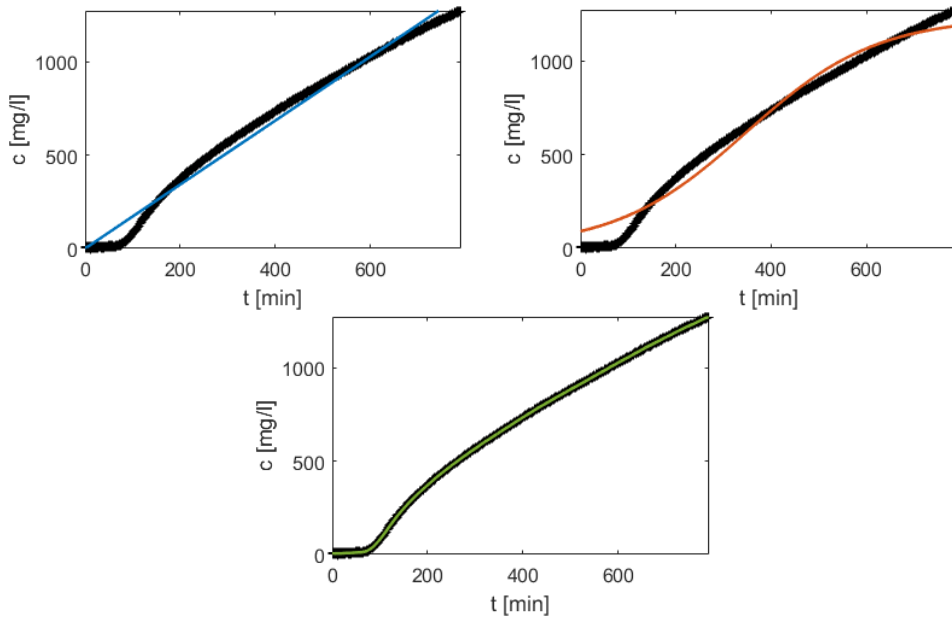
Fig. 3 All proposed estimates of the data set B data (black +), piecewise linear (blue), logistic (orange), kernel regression (green)

Since we have 24 different measurements of the concentration with respect to time for each data set, in Fig. 4 we can present boxplots of the resulting coefficients of determination, which describe the goodness-of-fit of the proposed models. The numerical results of the coefficient of determination, summarized in Tab. 1, show that the kernel method provides the best model of the concentration curve.
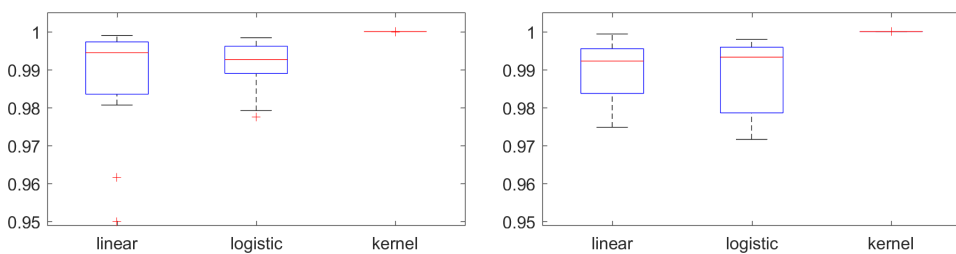

Fig. 4 Boxplots of the coefficients of determination for data A (left panel) and data B (right panel)

A general advantage of parametric models is their exact formula, which allows users to easily interpret the model and to simply calculate values of their interest. Namely, the piecewise linear model—a combination of two regression lines—provides a very simple tool for further usage.

Tab. 1 Boxplot summary statistics of the coefficient of determination

	Data A			Data B		
	linear	logistic	kernel	linear	logistic	kernel
maximum	0.9990	0.9984	1	0.9994	0.9979	1
upper quartile	0.9973	0.9961	1	0.9955	0.9959	1
median	0.9944	0.9926	1	0.9922	0.9933	1
lower quartile	0.9835	0.9890	1	0.9837	0.9786	1
minimum	0.7416	0.9775	0.9999	0.9748	0.9716	0.9999

Even though the kernel regression model does not have a compact equation, it has a great advantage; this model does not assume the shape of the estimated function beforehand. It is said that the nonparametric models let the data speak for themselves. That makes them more intuitive and flexible enough to grasp subtle aspects of the modelled data.

Focusing on disadvantages of the proposed methods, it needs to be pointed out that the piecewise linear model as well as the logistic model suffer from a prescribed shape of the function—what if the shape is wrongly assumed? The parametric methods are less robust than nonparametric ones. Although the nonparametric models, in which the suggested kernel smoothing belongs to, are more robust than the parametric models, their computing is slightly difficult. Also the prediction of these models is a more complicated task than the prediction in parametric settings.

In Tab. 2, the pros and cons of the proposed models are briefly summarized regarding their practical application.

Tab. 2 Pros and cons of the proposed smoothing models

Model	Pros (+)	Cons (−)
Linear	easy interpretation, quick evaluation	predefined shape, number of break points
Logistic	exact equation, easy interpretation	predefined shape, nonlinearity in parameters
Kernel	flexibility, data speak for themselves	bandwidth selection, no closed functional form

3.2 Lag-time Estimates

The next step in our analysis is to provide a procedure to estimate the lag-time value for which the model of the first derivative is needed. The parametric models have the straightforward calculation of their derivatives, see Eq. (3). The estimates of the first derivative of the kernel regression model were calculated using the central differences and the R functions *npreg*, *npregfast* and *npregderiv*.

In Figs 5 and 6, central differences of the raw data with the kernel regression estimates based on these differences (both in black) are displayed. In the lower panels of the figures, there are depicted the resulting estimates of the first derivatives using the approximation formula $\widehat{F'}(t) = b_1$, see Eq. (4) (in violet), and the direct estimate of the first derivative (in light blue).
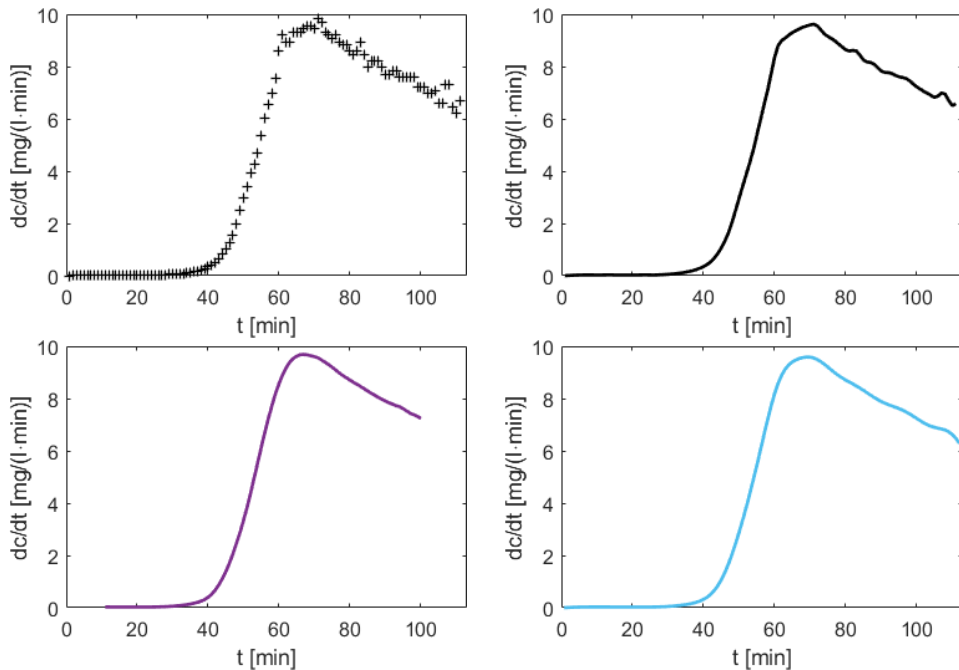
*Fig. 5 Estimates of the first derivative of data A central differences (data, black +),
kernel regression on central differences (kernel 1, black), estimate using
the approximation formula (kernel 2, violet) and the direct estimate (kernel 3, blue)*

All methods provide similar results of the first derivative estimate. Kernel regression based on the central differences (denoted as 'kernel 1') gives a slightly undersmoothed estimate, as we can see in Fig. 5. On the other hand, the direct estimate (kernel 3) is influenced by the so-called boundary effects, as we can see in Fig. 6. The estimate using the approximation formula (kernel 2) provides a shorter estimate, because it uses an additional smoothing parameter, which shows how many data points are used from each side of the estimation point [30]. Therefore, the whole estimated function is shorter by twice the size of this smoothing parameter.

Having obtained the estimates of the first derivative, we can continue with the lag-time calculations. We construct a tangent line to the concentration curve at its inflection point t_0 using the standard formula $c = c(t_0) + c'(t_0) \cdot (t - t_0)$. Then, we calculate its intersection with the time axis to obtain the lag-time value. In Tab. 3, the final equations of the tangent lines and their intersections with the time axis are summarized for data sets A and B.

From the graphs of the central differences (left upper panels of Figs 5 and 6), it is clear that the derivative estimates for data A will be more consistent than for data B.

Again, having 24 distinct measurements of the concentration with respect to time for each of the chemical test substance, we can compare the lag-time values obtained from our calculations with those estimated manually by an expert. Boxplots in Fig. 7 and their summary statistics in Tab. 4 provide a visual and numerical comparison of the lag-times.
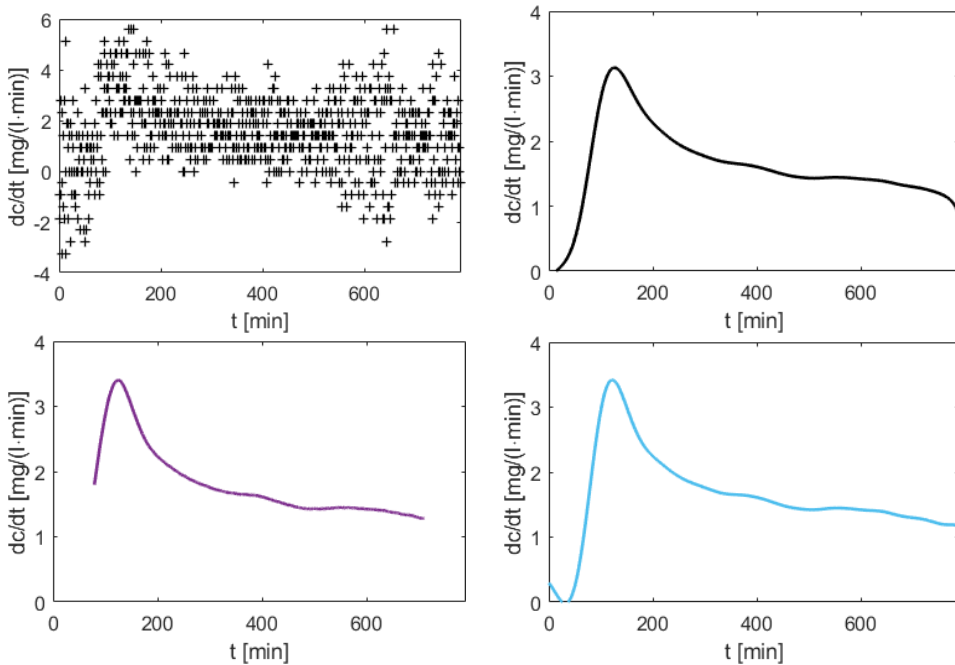
*Fig. 6 Estimates of the first derivative of data B central differences (data, black +),
kernel regression on central differences (kernel 1, black), estimate using
the approximation formula (kernel 2, violet) and the direct estimate (kernel 3, blue)*

From the boxplots and their values, it is obvious that only the linear model is different from other models. It is also supported by the analysis of variance with the post hoc Tukey's test, which shows that the only significantly different group is the lag-time estimated from the linear model. Moreover, we can see that the resulting kernel derivative estimates are close to the ones estimated manually by an expert.

The lag-time value reading is used to quickly and easily determine the value of the protection time of the barrier material used in field conditions. It should be noted that the evaluation of barrier properties in field conditions is different from the evaluation in laboratory conditions. For the purpose of rapid and relatively accurate evaluation of the

Tab. 3 Tangent lines and lag-time values for data sets A and B

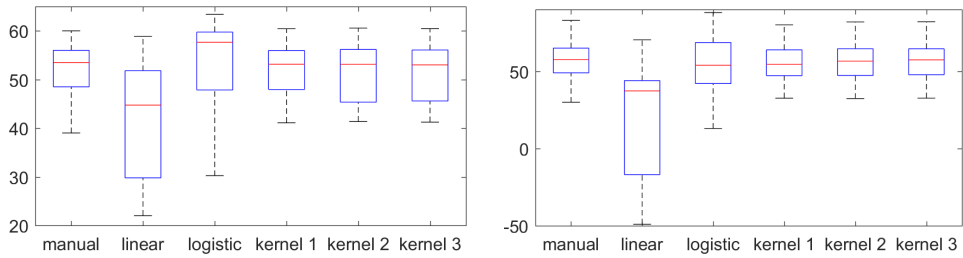| Model	Data A		Data B	
	tangent line	lag-time	tangent line	lag-time
linear	$c = 8.36t - 423.30$	50.66	$c = 1.72t + 0.18$	-0.10
logistic	$c = 10.85t - 620.77$	57.21	$c = 2.26t - 164.66$	72.95
kernel 1	$c = 9.62t - 511.45$	53.16	$c = 3.13t - 222.56$	71.10
kernel 2	$c = 9.69t - 515.55$	53.23	$c = 3.40t - 256.96$	75.51
kernel 3	$c = 9.59t - 509.05$	53.10	$c = 3.42t - 259.07$	75.78

Fig. 7 Boxplots of the lag-times [min] *for data A (left panel) and data B (right panel)*

Tab. 4 Boxplot summary statistics of the lag-time values [min]

Data A	manual	linear	logistic	kernel 1	kernel 2	kernel 3
maximum	60.0	58.9	63.4	60.4	60.6	60.5
upper quartile	56.0	51.8	59.8	56.0	56.2	56.1
median	53.5	44.7	57.6	53.1	53.1	53.0
lower quartile	48.5	29.8	47.8	48.0	45.3	45.6
minimum	39.0	−28.7	30.2	41.1	41.4	41.2
Data B	manual	linear	logistic	kernel 1	kernel 2	kernel 3
maximum	83.0	70.4	88.1	80.1	81.9	82.1
upper quartile	65.0	44.0	68.6	63.9	64.6	64.6
median	57.5	37.3	54.0	54.6	56.6	57.4
lower quartile	49.0	−16.9	42.1	47.2	47.3	47.8
minimum	30.0	−49.1	12.9	32.6	32.4	32.6

protective time of the evaluated barrier material, lag-time methods appear to be very reliable.

In addition, it appears that the difference between the values required by the respective standards makes no practical difference in cases where there is massive penetration of the test chemical through the barrier material under test in short times. In these cases, the lag-time readout method can be compared in terms of its factual relevance to methods based on achieving standard values for permeation masses and rates.

3.3 Outline of an Algorithm

From Fig. 7 and Tab. 4, it can be seen that the non-parametric estimates of the lag-time give similar results to those measured by an expert, i.e., those designated manually by clicking on the graph to enter a line that is tangent to the linear part of the graph.

As one of the objectives of this article was to provide a basis for an automation of the measuring procedure so that it is not necessary to determine values from the permeation curve manually, the steps of the proposed procedure for modelling the concentration and estimating the lag-time are summarized below.

1. Given the discretely measured concentration with respect to time, select the model and calculate the estimate of the concentration curve.
2. Calculate the first derivative of the concentration curve and find its maximum point (i.e., the inflection point of the concentration curve).

3. Construct the tangent line at the inflection point and find its intersection with the time axis, which produces the lag-time value.

As we can see, the kernel regression gives the best results in both phases of the estimation process. It fits the measured concentration the best from all proposed models (see the coefficients of determination in Fig. 4). Also, the resulting lag-time calculation has the smallest variance when compared to the parametric models (see Fig. 7). Although the nonparametric approach might seem more complicated than the parametric one, the consumption of the computer time is more or less the same as for the parametric models.

4   Conclusion

Discretely measured continuous data bring a task how to reconstruct the continuous function describing the variable of our interest. In this article, we presented a statistical method of reconstructing the concentration curve of the chemical substances with respect to time and proposed a mathematical procedure how to calculate the lag-time values. We focused on modelling the chemical substances which form ionic solutions in water, thus they are able to undergo dissociation.

The proposed smoothing methods can be applied to the substances with concentration determined from a different measured characteristics, such as methods based on the detection of changes in frequency, activity of the concentration of radioactive particles, Raman spectra, concentration of the substance based on the occurrence of specific peaks, etc.

As seen from the results, the linear regression was not a suitable tool for flexible modelling; therefore, it was convenient to employ more sophisticated methods, either parametric—represented here by the logistic model—or the nonparametric models. The kernel estimate was proposed as an alternative to the piecewise linear model and the logistic model. Since the article was inspired by real data sets, namely cyclohexylamine and pyridine, the presented models were tested on these data sets. The nonparametric model gave the best results of fitting the course of the concentration of the chemical substance over time. Also the estimates of the lag-time values based on the kernel estimates were the closest ones to the results obtained manually by an expert using the KONDUKTOTEST device.

Having a continuous model of the discretely measured data allows estimating other characteristics of the data themselves, such as a normalized break through time, which is connected to the concentration curve, or the lag-time, which is connected to the derivative of the concentration curve. The next step would be an automation of the measuring procedure, where one can, after the proposed steps of smoothing and calculating, read values of selected characteristics.

Acknowledgement

References

[1] OTRISAL, P., K. FRIESS, M. URBAN, S. BUNGAU, D.M. TIT, D.E. MOSTEANU, Z. MELICHARIK, C. BUNGAU and L. ALEYA. Barrier Properties of Anti-gas Military Garments, Considering Exposure to Gas Organic Compounds. *Science of the Total Environment*, 2020, **714**, 136819. DOI 10.1016/j.scitotenv.2020.136819.

[2] OTRISAL, P. Decontamination Modules Formed by the Czech Armed Forces Chemical Corps. *Croatian Journal of Education – Hrvatski Casopis za Odgoj i obrazovanje*, 2012, **14**, pp. 123-127.

[3] STEPANEK, B. and P. OTRISAL. The Development and Establishment Process of Centres of Excellence in North Atlantic Organization. *Croatian Journal of Education – Hrvatski Casopis za Odgoj i obrazovanje*, 2012, **14**, pp. 169-174.

[4] OTŘÍSAL, P. and S. HOŠKOVÁ-MAYEROVÁ. Selected Aspects of Barrier Materials Assessment as a Part of the Reaction on Threats and Risks Connected with CBRN Problems. In: *Decision Making in Social Sciences: Between Traditions and Innovations*. Cham: Springer, 2020, pp. 531-543. DOI 10.1007/978-3-030-30659-5_32.

[5] KELLNEROVÁ, E., K. BINKOVÁ and S. HOŠKOVÁ-MAYEROVÁ. Assessment of the Efficiency of Respiratory Protection Sevices against Lead Oxide Nanoparticles. In: *Models and Theories in Social Systems. Studies in Systems, Decision and Control 179*. Springer, 2019, pp. 257-272. DOI 10.1007/978-3-030-00084-4_14.

[6] OTRISAL, P., V. OBSEL, S. FLORUS, C. BUNGAU, L. ALEYA and S. BUNGAU. Protecting Emergency Workers and Armed Forces from Volatile Toxic Compounds: Applicability of Reversible Conductive Polymer-based Sensors in Barrier Materials. *Science of the Total Environment*, 2019, **694**. DOI 10.1016/j.scitotenv.2019.133736.

[7] OTRISAL, P., C. DIACONU, O. BRATU, F.I. RADU, Z. MELICHARIK and S.G. BUNGAU. New Approaches Regarding the Protection Forces' Health against the Effects of Some Toxic Substances. *Romanian Journal of Military Medicine*, 2019, **122**(3), pp. 104-107.

[8] TALHOFER, V. and S. HOŠKOVÁ-MAYEROVÁ. Method of Selecting a Decontamination Site Deployment for Chemical Accident Consequences Elimination: Application of Multi-Criterial Analysis. *ISPRS International Journal of Geo-Information*, 2019, **8**(4), 171. DOI 10.3390/ijgi8040171.

[9] PAL, T., G.D. GRIFFIN, G.H. MILLER, A.P. WATSON, M.L. DAUGHERTY and T. VO-DINH. Permeation Measurements of Chemical Agents Simulants through Protective Clothing Materials. *Journal of Hazardous Materials*, 1993, **33**, pp. 123-141. DOI 10.1016/0304-3894(93)85067-O.

[10] OTRISAL, P., V. OBSEL, J. BUK and L. SVORC. Preparation of Filtration Sorptive Materials from Nanofibers, Bicofibers, and Textile Adsorbents without Binders Employment. *Nanomaterials*, 2018, **8**(8), 564. DOI 10.3390/nano8080564.

[11] PRIKRYL, R., P. OTRISAL, V. OBSEL, L. SVORC, R. KARKALIC and J. BUK. Protective Properties of a Microstructure Composed of Barrier Nanostructured Organics and SiOx Layers Deposited on a Polymer Matrix. *Nanomaterials*, 2018, **8**(9), 679. DOI 10.3390/nano8090679.

[12] FLORUS, S., P. OTŘÍSAL and V. OBŠEL. *Methodology of KONDUKTOTEST for Quick Determination of Resistance of Porous (Filtration) and Non-porous (Insulation) Barrier Materials against the Static Permeation of Sulphur Mustard and Other Volatile Toxic Compounds with Acid-base Properties Dissolvable in Water on Ion Dilution* (in Czech). Vyškov: NBC Defence Institute, 2013.

[13] MATURO, F. and S. HOŠKOVÁ-MAYEROVÁ. Analyzing Research Impact via Functional Data Analysis: A Powerful Tool for Scholars, Insiders, and Research Organizations. In: *Proceedings of the 31st International Business Information Management Association Conference Innovation Management and Education Excellence through Vision 2020*. Milan: IBIMA, 2018, pp. 1832-1842.

[14] OTRISAL, P., Z. MELICHARIK, L. SVORC, S. BUNGAU, I. VIRCA, G. BARSAN and D. MOSTEANU. Testing Methods of Assessment for the Chemical Resistance of Insulating Materials Against the Effect of Selected Acids. *Revista de Materiale Plastice*, 2018, **55**(4), pp. 545-551. DOI 10.37358/MP.18.4.5071.

[15] RYMARCZYK, T., B. STEFANIAK, K. KANIA, M. MAJ and K. NIDERLA. Implementing Deterministic Methods to Solve the Inverse Problem for the Model with Lungs and Heart in the EIT. *Przeglad Elektrotechniczny*, 2020, **96**, pp. 128-128. DOI 10.15199/48.2020.07.23.

[16] VALIŠ, D., L. ŽÁK and Z. VINTR. Perspective Exploratory Methods for Multidimensional Data Analysis. In: *2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE, 2019, pp. 426-430. DOI 10.1109/IEEM44572.2019.8978643.

[17] OTŘÍSAL, P., S. FLORUS and R. KARKALIĆ. Resistance of Barrier Materials Against Toxic Compounds Permeation and Its Evaluation in Accordance with New European Norms. *International conference KNOWLEDGE-BASED ORGANIZATION*, 2017, **23**(3), pp. 224-233. DOI 10.1515/kbo-2017-0182.

[18] OTŘÍSAL, P. and S. FLORUS. Assessment of Chemical Resistance of Barrier Materials Based on Practical Application of Conductometry. *International conference KNOWLEDGE-BASED ORGANIZATION*, 2016, **22**(3), pp. 662-669. DOI 10.1515/kbo-2016-0114.

[19] MONTGOMERY, D.C., E.A. PECK and G.G. VINING. *Introduction to Linear Regression Analysis*. 4th ed. New Jersey: Wiley, 2006. ISBN 978-0-471-75495-4.

[20] MEYER, P.S., J.W. YUNG and J.H. AUSUBEL. Logistic Growth and Substitution: The Mathematics of the Loglet Lab Software Package [online], 2004 [viewed 2019-12-16]. Available from: https://phe.rockefeller.edu/LogletLab/whitepaper/LL1.html

[21] FAN, J. Local Linear Regression Smoothers and their Minimax Efficiencies. *The Annals of Statistics*, 1993, **21**, pp. 196-216. DOI 10.1214/aos/1176349022.

[22] WAND, M.P. and M.C. JONES. *Kernel Smoothing*. London: Chapman and Hall, 1995. ISBN 0-412-55270-1.

[23] HOROVÁ, I., J. KOLÁČEK and J. ZELINKA. *Kernel Smoothing in Matlab. Theory and Practice of Kernel Smoothing*. Singapore: World Scientific, 2012. ISBN 978-981-4405-48-5.

[24] SALHA, R. Kernel Estimation of the Regression Mode for Fixed Design Model. *Electronic Journal of Applied Statistical Analysis*, 2014, **7**(2), pp. 315-325. DOI 10.1285/i20705948v7n2p315.

[25] HAYFIELD, T. and J.S. RACINE. Nonparametric Econometrics: The np Package. *Journal of Statistical Software*, 2008, **27**(5), pp. 1-32. DOI 10.18637/jss.v027.i05.

[26] HURVICH, C.M., J.S. SIMONOFF and C.L. TSAI. Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society B*, 1998, **60**, pp. 271-293. DOI 10.1111/1467-9868.00125.

[27] RACINE, J.S. *Nonparametric Econometrics: A Primer*. Hanover (Mass.): Now Publishers, 2008. ISBN 978-1-60198-110-3.

[28] GASSER, T. and H.-G. MÜLLER. Estimating Regression Functions and Their Derivatives by the Kernel Method. *Scandinavian Journal of Statistics*, 1984, **11**(3), pp. 171-185. Available from: https://www.jstor.org/stable/4615954

[29] SESTELO, M., N.M. VILLANUEVA, L. MEIRA-MACHADO and J. ROCA-PARDIÑAS. npregfast: An R Package for Nonparametric Estimation and Inference in Life Sciences. *Journal of Statistical Software*, 2017, **82**(12), pp. 1-27. DOI 10.18637/jss.v082.i12.

[30] SAVCHUK, O. and A.A. VOLINSKY. Nonparametric Estimation of SiC Film Residual Stress from the Wafer Surface Profile. *Measurement*, 2021, **177**, 109238. DOI 10.1016/j.measurement.2021.109238.

[31] DE BRABANTER, K., J. DE BRABANTER, B. DE MOOR and I. GIJBELS. Derivative Estimation with Local Polynomial Fitting. *Journal of Machine Learning Research*, 2013, **14**, pp. 281-301.

[32] WANG, W.W. and L. LIN. Derivative Estimation Based on Difference Sequence via Locally Weighted Least Squares Regression. *Journal of Machine Learning Research*, 2015, **16**, pp. 2617-2641.

[33] SAVCHUK, O. and A. VOLINSKY. npregderiv: Nonparametric Estimation of the Derivatives of a Regression Function [online], 2020. R package version 1.0 [viewed 2021-10-13]. Available from: https://CRAN.R-project.org/package=npregderiv