Revised submission to IJCRE 2017-08-29 Accepted 2017-10-09 Bayesian Deconvolution of Vessel Residence Time Distribution

Thomas Huddle¹, Paul Langston^{2*}, Edward Lester¹

1. Advanced Materials Research Group; 2. Fluids & Thermal Engineering Group; Faculty of Engineering, University of Nottingham, Nottingham, NG7 2RD, United Kingdom

Abstract

Residence time distribution (RTD) within vessels is a critical aspect for the design and operation of continuous flow technologies, such as hydrothermal synthesis of nanomaterials (Cabanas, Darr et al. 2000). RTD affects product characteristics, such as particle size distribution. Tracer techniques allow measurement of RTD, but often cannot be used on an individual vessel in multiple vessel systems due to unsuitable exit flow conditions. However, RTD can be measured indirectly by removal of this vessel from the system and deconvoluting the resulting detected tracer profile from the original trace of the entire system.

This paper presents three models for deconvolution of RTD: BAY an application of the Lucy-Richardson iterative algorithm (Richardson 1972, Lucy 1974) using Bayes' Theorem, LSQ an adaptation of a least squares error approach (Blackburn 1970) and FFT a Fast Fourier Transform. These techniques do not require any assumption about the form of the RTD.

The three models are all accurate in theoretical tests with no simulated measurement error. For scenarios with simulated measurement error in the convoluted distribution, the FFT and BAY models are both very accurate. The LSQ model is the least suitable and the output is very noisy; smoothing functions can produce smooth curves, but the resulting RTD is less accurate than the other models.

In experimental tests the BAY and FFT models produce near identical results which are very accurate. Both models run quickly, but in real time control the runtime for BAY would have to be considered further. BAY does not require any filtering or smoothing here, and so potentially there are applications where it might be more useful than FFT.

Keywords: Residence Time Distribution; Deconvolution; Bayes' Theorem; Least Squares; FFT

* Corresponding author p.langston@nottingham.ac.uk +44 951 4177

1 Introduction

Levenspiel (Levenspiel 1999) describes deviations from ideal flow (plug or mixed) for process equipment such as reactors, heat exchangers, packed columns caused by channelling, recycling of flow, stagnant regions which can adversely affect performance, costs or product quality. Deconvolution of residence time distribution (RTD) from tracer signals flowing into and out of such equipment can be very useful in design optimisation and control.

Langston (Langston 2002) describes practical cases where the properties of a system are a complex mixture of several components. It is often useful to extract component compositions from system property measurements; such as in signal processing, coal blend analysis and measurement of particle size distribution. Vessels in series will also give rise to convoluted RTD's. Accurate deconvolution of individual vessel features is a challenging process.

An important application concerns residence time distribution measurements of reactors in continuous flow processes. In these systems the reactor is defined as the primary system volume in which the desired chemical transformation takes place. Often there will be more than one chemical reaction occurring at the required conditions, such as competing reactions or decomposition of the product. The various reaction pathways are dependent on the reaction parameters, which include residence time in the reactor; based on an understanding of the various pathway kinetics, optimal processing conditions usually requires a specific residence time with minimal distribution. Jumbam (Jumbam, Skilton et al. 2012) have demonstrated the ability to optimise the methylation of alcohols in continuous flow supercritical carbon dioxide; the yield and selectivity are shown to depend on residence time among other reaction parameters. Another practical example is the continuous hydrothermal synthesis of nanomaterials in supercritical fluids; Adschiri (Adschiri, Hakuta et al. 2001) first proposed a mechanism for the generation of metal oxides using this technique. With countless applications, the properties of nanomaterials are greatly influenced by particle size, which is often in turn dictated by residence time in the system reactor. For many nanomaterial species where nucleation is virtually instantaneous, particle growth can be directly related to time spent in the reactor, as studied by various groups including Norby (Norby, Jensen et al. 2013); thus the residence time distribution in the reactor will directly translate to particle size distribution in the product solution. The use of different reactor geometries by Blood (Blood, Denyer et al. 2004) for continuous hydrothermal nanomaterial synthesis have been investigated previously and it is well known that mixing dynamics has a significant impact on particle nucleation and growth. In addition to reactor design issues, other control variables for this continuous hydrothermal process were discussed by Lester (Lester, Aksomaityte et al. 2012), including temperature, flow rate, flow ratio, pipe diameter and pressure. A recent project scales this process from bench to pilot through to industrial scale, at over 100 tons per year capacity, with the FP7 Sustainable Hydrothermal Manufacturing of Nanomaterials (SHYMAN) project (Particles 2015, Shyman 2015). Better control of products from this process, and the inevitable changes in product quality (as a result of scaling the process) will become a focus for the industry and therefore understanding residence time distributions and mixing dynamics is essential.

Spectral analysis methods are widely used for extracting components of signals (Blackburn 1970). One method uses a least squares technique, which sets up equations defining the sum of the error (predicted – actual signal) squares as a function of the predicted composition. This is differentiated with respect to the composition, set equal to zero and so the "best" composition

found. Simmons (Simmons, Langston et al. 1999) used a similar technique to predict particle size distribution from chords that was partially successful.

Least squares methods were used in the following RTD analysis. Kuu (Kuu 1992) used a numerical deconvolution method with Powell's non-linear least-squares algorithm to search for values of dispersion coefficients of solute molecules in a drug delivery system. Bruce (Bruce, Sai et al. 2004) experimentally measured the RTD in a turbulent bed contactor in series with an ideal Mixed Flow Tank. The RTD of the ideal mixed flow tank was deconvoluted from the RTD of the entire system to obtain the RTD of turbulent bed contactor alone as a function of Peclet number and mean residence time of an axial dispersion model; optimal values were found using a Matlab-based least squares function. Boskovic (Boskovic and Loebbecke 2008) modelled RTD in micromixers; deconvolution involved minimization of sum of error square by the unconstrained quasi-Newton method; two flow models were developed, one based on axial dispersion, the other based on statistical distributions allowing a certain skewness; the latter gave a better fit to the range of their experiments. Adeosun (Adeosun and Lawal 2009) describes numerical and experimental mixing studies in a flow micromixer and used the convolution technique, in which time-domain curve fitting via non-linear optimization was used to estimate parameters in the RTD model.

The Fast Fourier Transform FFT has been used in RTD analysis. Viitanen (Viitanen 1997) applied a FFT to deconvolute RTDs of process equipment with examples from aquifers in bedrock and industrial equipment with recirculation; this method allows filtering in the frequency domain of data with poor statistics. Gooseff (Gooseff, Benson et al. 2011) estimated RTDs in stream surface transient storage zones via signal deconvolution with Fourier and Laplace transforms.

A method to estimate RTD in complex systems is to use a Bayesian approach with a Markov Chain Monte Carlo model to infer presumed RTD form parameters. Krone-Davis (Krone-Davis, Watson et al. 2013) analysed pesticide reduction in constructed wetlands using a tanksin-series model. Massoudieh (Massoudieh, Leray et al. 2014) characterised groundwater flow systems using transient environmental tracer data.

Langston (Langston, Burbidge et al. 2001) used Bayes' theorem iteratively to predict diameter distributions of spherical particles or droplets from measurements of chords. It was then further developed (Langston and Jones 2001) to analyse non-spherical shapes in 2D; (Langston 2002) developed the model for general cases of component mixture deconvolution and theoretically compared it to a least squares method. The latter was shown to be generally more applicable and robust, but the Bayes' model was, in some cases, less susceptible to simulated noise. Subsequent to publication of these papers the authors found a similar method in the Lucy-Richardson algorithm (Richardson 1972, Lucy 1974) developed for image analysis.

The literature review can be summarised as follows. Measurement, modelling, prediction and control of RTD is generally useful in reactor design and operation. Bayesian Markov Chain Monte Carlo methods are useful in complex systems to infer RTD from measurements. Deconvolution of RTD is challenging and useful. Some deconvolution methods assume a form of RTD and some do not; it is assumed here that the latter is potentially more accurate. Least squares and FFT are established methods. To the authors' knowledge an iteration on Bayes' theorem such as the Lucy-Richardson algorithm has not been used for RTD deconvolution and there is little comparison of the accuracy and speed of these different approaches.

In this paper, Section 2 describes the development of three models to deconvolute vessel RTDs: an iterative model based on the Lucy-Richardson algorithm, a least squares model based on Blackburn (Blackburn 1970), and a Fast Fourier Transform as in (Viitanen 1997); in these models no assumption is necessary about the form of the RTD. Section 3 applies the models in Matlab to a theoretical case to directly test theoretical accuracy; this is initially tested without including any replicated error, then with replicating measurement error; Section 3 goes on to describe an experimental study to test these models. Section 4 presents the conclusions.

2 Models of RTD Deconvolution

2.1 Vessel RTD Convolution

Levenspiel (Levenspiel 1999) describes how the residence time distribution of a fluid passing through a vessel can be normalised and represented by E the exit age distribution for a closed vessel boundary condition (fluid only enters and leaves vessel once). A one-shot tracer signal entering (C_{in}) and leaving a vessel (C_{out}) can be represented by the convolution integral

$$C_{out}(t) = \int_{0}^{t} C_{in}(t') E(t-t') dt$$

$$C_{out} = E C_{in}$$

If we have three independent flow units a, b, c, closed and connected in series then

$$C_{out} = E_a E_b E_c C_{in}$$

Levenspiel (Levenspiel 1999) describes that independence implies that a fluid loses its memory between vessels and if an element flows faster in one vessel it is no more likely to flow faster in the next vessel, which is not always the case with laminar flow.

Section 2 now describes how deconvolution of RTD can be undertaken using either:

- Bayes' Theorem model referred to later as BAY
- Least Squares LSQ
- Fast Fourier Transform FFT
- Polynomial Division.

2.2 RTD Deconvolution using Lucy-Richardson algorithm model BAY

A far more challenging exercise is deconvolution; ie if C_{out} , C_{in} , E_a , E_b , are known then calculate E_c . This can be more generally represented by a two-vessel system: say E_b is unknown, E_a and E_T are known,

$$E_T = E_a E_b$$

Lee (Lee 2012) describes Bayes' theorem as follows. If $A_1, ..., A_m$ are a set of mutually exclusive, exhaustive events in a possibility space S_p , and B is any other event in S_p such that P(B)>0, then

$$P(A_j | B) = \frac{P(B | A_j)P(A_j)}{\sum_{k=1}^{m} P(B | A_k)P(A_k)} \quad j = 1, 2, ..., m$$

This is applied here using the iterative procedure used in image analysis (Richardson 1972, Lucy 1974) and modified for estimating diameter distributions form chord measurements (Langston, Burbidge et al. 2001).

Dividing *E* into *N* equal time bins *j*:

- 1 Assume a uniform distribution in $E_{bj} = 1.0/N$, for all *j* (or input an initial estimate)
- 2 For each *j* use Bayes' theorem to calculate probability P_{ji} equivalent to E_{bjTi} , which can be considered as probability that a fluid element which exits total system at time bin *i* resided in vessel b for *j* time bins

$$E_{bjTi} = \frac{E_{Tibj}E_{bj}}{\sum_{k=1}^{N} E_{Tibk}E_{bk}} \text{ for each } i$$
$$E_{Tibj} = E_a(i - j + 1)$$

3 For each *j* recalculate (note denominator equals unity)

$$E_{bj} = \frac{\sum_{i=1}^{N} E_{bjTi} E_{Ti}}{\sum_{i=1}^{N} E_{Ti}}$$

4 If E_b has changed significantly repeat from step 2 using the new values of E_b ; however, in this paper the model is run for a specified number of iterations, n_{itr} .

It is noted here that E is being used in a discretized form representing bin probabilities rather than a continuous probability density (hence the +1 in step 2 above); this is also applicable to the RTD plots shown later. As indicated above, for Bayes' theorem to be applicable the flow units must be independent. The time interval (bin size) must also be small relative to the overall time.

2.3 RTD Deconvolution using Least Squares model LSQ

Blackburn (Blackburn 1970) describes the problem in spectral analysis as follows:

$$\sum_{j=1}^{m} X_{j} R_{ij} + e_{i} = S_{i} \qquad i = 1, \dots, n$$

where

 S_i is the amplitude of the data spectrum at the ith channel or sample point R_{ij} is the amplitude of the jth reference spectrum at the ith channel X_j is the set of coefficients to be extracted e_i is the noise or error spectrum n is the number of channels m is the number of standards.

Summing over the *n* channels obtains δ the sum of the error squares. A least-square error approach based on calculating $(d\delta/dX_k)$ for k = 1, ..., m, gives a set of *m* linear equations of the form:

$$\sum_{i=1}^{n} S_{i} R_{ik} = \left[\sum_{j=1}^{m} \left[\sum_{i=1}^{n} R_{ij} R_{ik} \right] \right] X_{j}$$

These can be solved to give optimal X_j values. This requires a matrix R_{ij} , but we have a vector for each vessel RTD. Hence the model is applied here to a two vessel deconvolution by formulating R_{ij} from the vector E_a :

 $R_{ij} = E_{ak}$, where k=i-j+1 and k > 0 and k < N+1, else $R_{ij} = 0$

where *i* is the subscript of E_T (=*S* above) and *j* is the subscript of E_b (=*X* above).

The output from this model is generally very noisy here. Integral smoothing functions available in Matlab were investigated. General smoothing algorithms such as the default moving average or Savitzky-Golay filters, are not well suited, regardless of the polynomial degree specified, as the noise on the response data appears to be equal across all time bins (including where the probability of E_b is actually zero). Smoothing which operate on localised ranges of time bins appear more useful. The most accurate was found to be a local regression method using weighted linear least squares and a 2nd degree polynomial model (*'loess'* in Matlab); this method is implemented using the Matlab *smooth* function with the *'loess'* method, with optimal values for the *span* parameter varying between 0.2 and 0.5 and the function is applied twice in a loop; additionally, any negative values for the LSQ result are set to zero, although this makes little difference to the overall accuracy.

2.4 RTD Deconvolution using Fast Fourier Transform model FFT

As noted in section 1 Viitanen (Viitanen 1997) applies FFT to deconvolute RTDs of process equipment. In the frequency domain w the convolution becomes a straightforward multiplication and the deconvolution a division.

$$E_T(w) = E_a(w) \cdot E_b(w)$$
$$E_b(w) = E_T(w) / E_a(w)$$

 $E_b(t)$ can then be obtained by the inverse FFT after appropriate filtering in the frequency domain to account for measurement error. This method is applied here using Matlab functions. The filtering at high and low frequencies (see code on-line) is essential here. Also truncating any values below zero improved the performance of the model here.

2.5 RTD Deconvolution using Polynomial Division

RTD deconvolution is similar in principle to a more general vector deconvolution. The function in Matlab *deconv* does this by polynomial division. This has been tested here for RTD deconvolution; it is accurate for theoretical cases but is highly inaccurate in practice with measurement errors; in this study the method even failed to produce meaningful results with practical or simulated measurement error. This reason is clear on considering the principle of the method where errors in the coefficients of the higher terms are propagated in sequential division. Hence it is concluded that it is not worth considering this method further here.

2.6 Measures of Accuracy

To assess the models the deconvoluted RTD is compared with the known RTD using three methods:

- Mean Square Error *M*_{SE}
- Coefficient of Determination *R*²
- Overlapping Coefficient O_{VL} (Clemons and Bradley 2000)

$$M_{SE} = \frac{1}{N} \sum_{i=1}^{N} (E_i - E_{bi})^2$$

Where E_i is the estimated probability for each bin, E_{bi} is the known target probability for each corresponding bin, and N is the total number of bins.

$$R^{2} = 1 - \frac{\sum_{i=1}^{N} (E_{i} - E_{bi})^{2}}{\sum_{i=1}^{N} (E_{bi} - \overline{E_{bi}})^{2}}$$
$$O_{VL} = \sum_{i=1}^{N} \min[E_{i}, \quad E_{bi}]$$

3 Application of RTD Deconvolution

3.1 Theoretical Examples

3.1.1 Theoretical case with no Simulated Measurement Error

Levenspiel (Levenspiel 1999) describes types of flow models: the dispersion model and tanksin-series model which are similar and widely applicable, and the pure convection model for laminar flow in short pipes. For simplicity initial theoretical tests were undertaken on deconvolution of two normal distributions with no simulated measurement error. All the models, BAY, LSQ and FFT are accurate in these tests.

In order to investigate the potential accuracy of the models when tackling the deconvolution of non-normal convoluted data, as found in our reactor systems, two theoretical "shouldered" distributions were generated through the additive combination of normal distributions with differing parameters. These simulated examples more closely represent practical distributions generated for tubular reactor residence times, where tailing of the peaks is a common feature.

In generating the shouldered distributions, the underlying normal distributions were subjected to a vertical scaling factor in order to control their influence on the combined distribution; the resultant shouldered distributions were then converted to probability distributions through division by their total integral. Distribution A was generated by the additive combination of two normal distributions, and distribution B from three normal distributions using the parameters indicated in Table 1.

	Mean (µ)	Standard Deviation (σ)	Vertical scaling	Bins
Distribution A				100
Normal 1	20.0	2.5	1.00	
Normal 2	27.5	3.0	0.75	
Distribution B				100
Normal 1	30.0	5.0	0.50	
Normal 2	42.5	5.0	0.35	
Normal 3	55.0	5.0	0.20	

Table 1: Parameters used to produce realistic distributions A and B; then converted into bin probabilities E_a and E_b .

The probability distributions for A and B (E_a and E_b) are plotted along with that of the convolution of A and B (E_t) in Figure 1 with a time bin size of 1s. Figure 2 illustrates successive iterations of the BAY model in deconvoluting E_a from E_T in order to generate an estimate of E_b ; the true value of E_b and the estimate from the model are graphically indistinguishable after 100 iterations. Table 2 shows a summary of the statistical measures of fit for the models described in section 2. Clearly all three models are accurate here, with FFT the most accurate.



Figure 1: Input probability distributions E_a and E_b produced from theoretical nonnormal distributions, and their convoluted product E_T



Figure 2: Plot of the Bayes iterative approximations of E_b , represented as E_n ; curves for E_b and for the BAY $n_{itr} = 100$ estimate are graphically colinear in this theoretical no error case

	M_{SE}	R^2	O_{VL}
BAY $n_{itr} = 10$	8.4e-7	0.9949	0.9748
BAY <i>n</i> _{itr} =100	3.8e-9	1	0.9984
BAY <i>n</i> _{itr} =1000	4.1e-11	1	0.9998
LSQ	7.3e-7	0.9956	0.9709
FFT	9.2e-13	1	1

 Table 2: Model accuracy in calculation of RTD of B for theoretical case with no replicated measurement error

3.1.2 Case of non-Normal Distribution with Simulated Measurement Error

In practical scenarios, experimental inaccuracy and bias may produce error in any measured residence time. In order to evaluate the effect on the deconvolution method, simulated error has been applied to the convoluted distribution generated using the parameters detailed in section 3.1.1. The probability distribution E_a can then be deconvoluted from the erroneous distribution E_{T-err} in order to observe the effect on the estimation of E_b .

Several types of error are observed in practice: a bias, i.e. shifting the distribution by a fixed time; noise in the distribution, resultant of detector issues (when measuring residence time by tracer detection); scaling of the distribution, i.e. compression or expansion of the distribution resulting in a similar shape, but occurring over a different range of time (pump flow rate inaccuracies/inconsistencies).

Unpublished work within the Lester group involving the injection and detection of tracers has shown the last of these error types (scaling) to be most frequently observed. Figure 3 illustrates an example of a simulated "scaling error" applied to E_T , with 10% compression scaling applied about the bin at which the probability of E_T is at maximum, in order to generate E_{T-err} . The comparisons of the true B and estimated B for the models is shown in Figure 4 and the statistical measures of accuracy are shown in Table 3.



Figure 3: Input probability distributions, with time bin of 1s, E_a and E_b used to produce E_T . E_T has then been subjected to compression scaling by 10 % about the bin of its maximum probability to generate E_{T-err} . This is reasonably similar to frequently observed experimental error.

Table 3: Model accuracy in calculation of RTD of B (E_b) for theoretical case with simulated 10% compression of AB (E_{T-err}) to replicate measurement error

	$M_{SE}{ t x10^{-6}}$	R^2	O_{VL}
BAY $n_{itr} = 10$	5.79	0.9648	0.9267
BAY <i>n</i> _{itr} =100	5.36	0.9673	0.9314
BAY <i>n</i> _{itr} =1000	5.34	0.9674	0.9315
LSQ	16.3	0.9007	0.8736
FFT	5.04	0.9692	0.9346

SEE OVERPAGE FOR FIGURE

Figure 4: Plots comparing BAY n_{itr} =100, LSQ and FFT methods for approximating B from theoretical deconvolution of A from total AB using replicated measurement error of AB



The FFT and BAY models are both very accurate, with FFT marginally the best. 10 iterations for the BAY model seems sufficient for practical purposes. The LSQ model is the least suitable; Matlab produces a warning message on solution of the matrix equation *Matrix is close to singular or badly scaled*, and the output is very noisy; smoothing functions can produce smooth curves, but the resulting RTD is less accurate than the other two models. (Further investigation for zero replicated error cases, not shown here, indicates that LSQ is accurate, not noisy and no smoothing is required for about 40 or fewer time bins in E_a and in E_b .) Further investigation is required to test if this method could be improved, it is not considered further in this paper.

3.2 Experimental Examples

3.2.1 Experimental Method



Figure 5: Diagram showing the apparatus configuration used for measuring residence time distributions inside volumes A and B. The "inj" volume can be considered as the above system without A and B present.

All the experiments described here were undertaken by the authors producing all the experimental results in section 3. A schematic of the experimental configuration is illustrated in Figure 5. A Gilson 305 pump was used to deliver a flow of deionised water through the system, with the flow rate controlled by computer script (via a Gilson 506C interface not shown in illustration). The flow was directed through an electronically actuated Rheodyne 6 port 2-position switching valve mounted on a Gilson 234 autosampler (not shown in illustration). In the "load" position, the switching valve allowed water from the pump to flow through the valve and into the system (this valve position is indicated in the illustration); the autosampler was programmed to fill the 10 μ l sample loop of the switching valve with an aqueous solution containing 10g.l⁻¹ Nigrosin dye (Fischer Scientific) to be used as a tracer.

When instructed by the computer script, the valve switches to the "inject" position, which would introduce the fixed volume of tracer solution into the system at the predetermined flow rates. After passing through the system volumes of interest, the concentration of tracer detected at the system outlet was continuously measured using a Gilson 119 UV/vis detector, set to monitor at a wavelength of 545 nm (visible).

The system volume was divided into two sections, referred to as sections A and B. The primary volume of section A was a 200 mm length of stainless steel tubing (0.25" outer diameter, 2.34 ml volume) and section B a 300 mm length of the same tubing (3.51 ml volume). Both volumes also included short sections of 0.0625" tubing, which allowed sections A or B to also be investigated independently, with no requirement for additional tubing/fittings in any configuration.

Four flow rates were investigated: 10, 15, 20 and 25 ml.min⁻¹. For each flow rate, distribution measurements were acquired for volume A, volume B and both volumes in series (referred to as volume AB), as shown in Figure 6. The intention was to use the deconvolution models to estimate volume B by deconvolution of measured volume A from the measured volume AB. The estimated volume B could then be compared against the measured volume B in order to assess model accuracy in a practical scenario. Initial experiments pointed to a systematic time delay between estimated B and measured B whenever A was deconvoluted from AB. This timing discrepancy was seen to be identical when deconvoluting B from AB (in comparing estimated A with measured A). This discrepancy was attributed to both the additional volumes both before and after volume AB (i.e. connection to the sample loop and internal volumes of the UV/vis detector) which could not be removed from the system, in addition to a timing delay between the valve switching and the start of UV/vis absorption data acquisition. In order to account for this delay, the additional volume was measured separately using the same procedure, essentially measuring the system volume in absence of A and B (this measurement is referred to as "inj"). The timing of the resultant peak was found to exactly account for the timing discrepancy observed in deconvolution of either A or B. Figure 6 shows experimentally acquired distributions for A, B, AB and "inj". For distributions A, B and AB, the time bins at which absorption is first detected are 124, 205 and 292 respectively. For independently measured "inj", absorption is first detected at bin 40; this accurately accounts for the discrepancy observed between first detection of AB (290) and the combination of A and B (124+205).



Figure 6: Example UV-Vis absorption traces against time for volumes A, B, A and B combined and the additional volume-time delay "inj". The timing of distribution "inj" accounts for the discrepancy observed when convoluting measured A and B and comparing to measured AB

When A, B and AB are measured separately, this timing delay and additional volume is included once per measurement. Therefore when the measured A or B distribution is deconvoluted from the measured AB distribution, the resulting estimation is that for A or B without the timing delay or additional volume. Although this is the intention in a practical scenario of residence time distribution measurements, the equivalent physical measurement cannot be obtained for comparison and evaluation of error in the deconvolution process. Rather than multiple deconvolution steps (which would compound any error in the deconvolution process), the estimation of distribution B can be convoluted with the measurement of "inj" as indicated in Figure 7; this allows for direct comparison with the measured distribution of B and to assess the accuracy of a single implementation of the deconvolution methods.

Step 1: Deconvolution



Figure 7: Schematic showing the steps taken in order to assess the accuracy of a single implementation of the deconvolution methods. After deconvoluting measured A from measured AB, the resulting approximation must be convoluted with the measurement of "inj" (at the specific flow rates) in order to be comparable with individually measured B

Although the intention is to apply the deconvolution methods to a scenario with genuine error, efforts were made here to minimise experimental error where possible. Experimental measurements of flow distributions at each flow rate were therefore performed 5 times for each configuration (inj, A, B and AB). A hypothetical mean trace was created for each set of measurements, and the closest of the measurements (based on minimum mean square error) to the hypothetical mean was selected. This ensured that "real" data was used, but that the distribution was a fair representation rather than an outlying result. Figure 8 shows an example case of the measured RTDs.



Figure 8: Experimental tracers as probability distributions for flow 15 ml.min⁻¹ showing RTD for: individual volumes A, B; and for the two vessels in series with an extra "inj" volume to aid illustration of experimental error; and for A and B convoluted (which then counts "inj" twice). The comparison of the two AB curves is effectively an indication of experimental error.

3.2.2 Comparing model and experimentally measured results

The measured distribution of A was deconvoluted from that of AB to generate an estimation of B at four different flowrates. The deconvolution was conducted with models BAY and FFT since these looked the most promising in section 2. The predicted distributions and actual distribution are plotted in Figure 9; the model accuracy is shown in Tables 4 and 5 along with example CPU computer runtimes.



Figure 9: Plots comparing FFT and BAY n_{itr} =100 methods for approximating B from experimentally measured distributions of AB and A for flowrate 15 ml.min⁻¹. (blue dot: experimentally measured points; red line: model)

Flow rate (ml.min ⁻¹)	M_{SE} x10 ⁻⁸		R^2		O_{VL}	
	FFT	BAY	FFT	BAY	FFT	BAY
10	1.04	1.04	0.9929	0.9929	0.9461	0.9462
15	5.72	5.72	0.9799	0.9798	0.9146	0.9146
20	7.09	7.08	0.9840	0.9841	0.9148	0.9147
25	2.40	2.40	0.9961	0.9961	0.9554	0.9554

Table 4: Accuracy of model fit to experimental RTD of B for FFT and BAY $n_{itr} = 100$

Table 5: Accuracy of model fit to experimental	RTD of	f B for	flowrate	10 r	nl.min ⁻¹	for
FFT and BAY <i>n_{itr}</i> variation with CPU times						

	M_{SE} x10 ⁻⁸	R^2	O_{VL}	CPU (s)
BAY $n_{itr} = 5$	3.20	0.9779	0.9268	4.3
BAY $n_{itr} = 10$	1.13	0.9923	0.9415	7.1
BAY $n_{itr} = 100$	1.04	0.9929	0.9462	65
FFT	1.04	0.9929	0.9461	0.4

Essentially the BAY and FFT models produce near identical estimates which are very accurate in all four cases. There is some variation in experimental error, which in all cases is less than the theoretical error introduced in section 3.1.2.

The CPU times should be considered as a guide only. They are undertaken with Matlab R2015a on a high spec desktop PC, 3.5GHz Intel processor. There is some small variation with repeated simulations. These are simulations without plotting the figures since this output takes a significant time in proportion. Note that FFT uses library routines but BAY runs mainly with author code, which would be significantly faster if compiled. The runtimes are short and this would only be an issue in real time control systems with many repeated simulations. If BAY were to be used in such a situation this issue would need further consideration.

4 Conclusions

The three RTD deconvolution models, BAY, LSQ, FFT are all accurate in theoretical tests with no simulated error.

Further tests incorporating simulated error in the convoluted distribution showed that FFT and BAY are very accurate; LSQ is the least suitable.

The polynomial division method is fundamentally unsuitable.

Experimentally tested BAY and FFT models produced near identical results which are very accurate.

 R^2 and O_{VL} are useful statistical measures here.

BAY does not require any filtering or smoothing here, and so potentially there are applications where it might be practically the most useful, although there may be a runtime penalty.

Acknowledgements

This work was funded through the European Union's Seventh Framework Programme (FP7/2007–2013), grant agreement no. FP7-NMP4-LA-2012-280983, the SHYMAN project. Anonymous advice on a previous draft regarding use of O_{VL} and FFT has been incorporated.

Further information can be accessed on: *Journal website URL*. This comprises the Matlab code for section 3.1.4, the theoretical assessment with replicated measurement error.

References

Adeosun, J. T. and A. Lawal (2009). "Numerical and experimental mixing studies in a MEMSbased multilaminated/elongational flow micromixer." <u>Sens. Actuators, B</u> **139**(2): 637-647. Adschiri, T., Y. Hakuta, K. Sue and K. Arai (2001). "Hydrothermal synthesis of metal oxide

nanoparticles at supercritical conditions." <u>J. Nanopart. Res.</u> **3**(2-3): 227-235.

Blackburn, J. A. (1970). Spectral analysis: methods and techniques. New York,, M. Dekker.

Blood, P. J., J. P. Denyer, B. J. Azzopardi, M. Poliakoff and E. Lester (2004). "A versatile flow visualisation technique for quantifying mixing in a binary system: application to continuous supercritical water hydrothermal synthesis (SWHS)." <u>Chem. Eng. Sci.</u> **59**(14): 2853-2861.

Boskovic, D. and S. Loebbecke (2008). "Modelling of the residence time distribution in micromixers." <u>Chemical Engineering Journal</u> **135**: S138-S146.

Bruce, A. E. R., P. S. T. Sai and K. Krishnaiah (2004). "Characterization of liquid phase mixing in turbulent bed contactor through RTD studies." <u>Chemical Engineering Journal</u> **104**(1-3): 19-26.

Cabanas, A., J. A. Darr, E. Lester and M. Poliakoff (2000). "A continuous and clean one-step synthesis of nano-particulate Ce1-xZrxO2 solid solutions in near-critical water." <u>Chemical Communications(11)</u>: 901-902.

Clemons, T. E. and E. L. Bradley (2000). "A nonparametric measure of the overlapping coefficient." <u>Computational Statistics & Data Analysis</u> **34**(1): 51-61.

Gooseff, M. N., D. A. Benson, M. A. Briggs, M. Weaver, W. Wollheim, B. Peterson and C. S. Hopkinson (2011). "Residence time distributions in surface transient storage zones in streams: Estimation via signal deconvolution." <u>Water Resources Research</u> **47**.

Jumbam, D. N., R. A. Skilton, A. J. Parrott, R. A. Bourne and M. Poliakoff (2012). "The Effect of Self-Optimisation Targets on the Methylation of Alcohols Using Dimethyl Carbonate in Supercritical CO2." J. Flow. Chem. 2(1): 24-27.

Krone-Davis, P., F. Watson, M. Los Huertos and K. Starner (2013). "Assessing pesticide reduction in constructed wetlands using a tanks-in-series model within a Bayesian framework." <u>Ecol. Eng.</u> **57**: 342-352.

Kuu, W. Y. (1992). "Determination of Residence-Time Distribution in Iv Tubing of in-Line Drug Delivery System Using Deconvolution Technique." <u>International Journal of Pharmaceutics</u> **88**(1-3): 369-378.

Langston, P. A. (2002). "Comparison of least-squares method and Bayes' theorem for deconvolution of mixture composition." <u>Chemical Engineering Science</u> **57**(13): 2371-2379.

Langston, P. A., A. S. Burbidge, T. F. Jones and M. J. H. Simmons (2001). "Particle and droplet size analysis from chord measurements using Bayes' theorem." <u>Powder Technology</u> **116**(1): 33-42.

Langston, P. A. and T. F. Jones (2001). "Non-spherical 2-dimensional particle size analysis from chord measurements using Bayes' theorem." <u>Part. Part. Syst. Char.</u> **18**(1): 12-21.

Lee, P. M. (2012). <u>Bayesian statistics : an introduction</u>. Chichester, West Sussex ; Hoboken, N.J., Wiley.

Lester, E., G. Aksomaityte, J. Li, S. Gomez, J. Gonzalez-Gonzalez and M. Poliakoff (2012). "Controlled continuous hydrothermal synthesis of cobalt oxide (Co3O4) nanoparticles." <u>Prog.</u> <u>Cryst. Growth Charact. Mater.</u> **58**(1): 3-13.

Levenspiel, O. (1999). Chemical reaction engineering. New York, Wiley.

Lucy, L. B. (1974). "Iterative Technique for Rectification of Observed Distributions." <u>Astronomical Journal</u> **79**(6): 745-754.

Massoudieh, A., S. Leray and J. R. de Dreuzy (2014). "Assessment of the value of groundwater age time-series for characterizing complex steady-state flow systems using a Bayesian approach." <u>Appl. Geochem.</u> **50**: 240-251.

Norby, P., K. M. O. Jensen, N. Lock, M. Christensen and B. B. Iversen (2013). "In situ synchrotron powder X-ray diffraction study of formation and growth of yttrium and ytterbium aluminum garnet nanoparticles in sub- and supercritical water." <u>R. Soc. Chem. Adv.</u> **3**(35): 15368-15374.

Particles, P. (2015). "Promethean Particles." from <u>www.prometheanparticles.com</u>.

Richardson, W. H. (1972). "Bayesian-Based Iterative Method of Image Restoration." Journal of the Optical Society of America **62**(1): 55-+.

Shyman. (2015). "Shyman." from <u>www.shyman.eu</u>.

Simmons, M. J. H., P. A. Langston and A. S. Burbidge (1999). "Particle and droplet size analysis from chord distributions." <u>Powder Technology</u> **102**(1): 75-83.

Viitanen, P. (1997). "Experiences on fast Fourier transform as a deconvolution technique in determination of process equipment residence time distribution." <u>Applied Radiation and Isotopes</u> **48**(7): 893-898.