NONPARAMETRIC AND SEMIPARAMETRIC MODELS IN COMPARISON OF OBSERVATIONS OF A PARTICLE-SIZE DISTRIBUTION

Walter Liggett*

ABSTRACT

Testing hypotheses about pairs of unnormalized histograms motivates this paper. The histograms contain particle counts for particle-size intervals. The analysis involves generalized-linear-model fitting of cubic splines with irregularly-spaced knots. Of interest is testing the null hypothesis that two sets of particle counts correspond to intensity functions that differ only by a scale factor and a constant shift in horizontal registration. An unknown smooth function is common to the two intensities. The alternative hypothesis is that in addition, the difference between the two intensities is also an unknown smooth function. We consider three approaches to knot placement. First is specification of so many knots that adequate representations of the unknown functions cannot be doubted. Second is data-driven choice of knots. Third is choice of knots based on prior knowledge of what intensity differences are plausible. For the data at hand, we show that specification of too many knots leads to tests with too little power and that data-driven knot selection can lead to false rejection of the null hypothesis. The data at hand seem to call for use of prior knowledge to construct a semiparametric model that incorporates the distinction between the two hypotheses in the parametric part.

1. Introduction

Consider a material property characterized not by a few parameters but by a smooth function $\lambda$ of an independent variable $r$, and consider measurement of such a property. A measurement consists of a collection of function values $\lambda(r_i)$ observed with noise. Replication of such a measurement involves repeating the entire measurement process on a sample of material with the same smooth function. Thus, a model of two or more replicate measurements includes at least one unknown smooth function. Measurement mechanisms known to cause differences between replicate measurements often have effects that can be described by a parametric linear model. In this case, the model of replicate measurements will be a partial linear model, a semiparametric model (Severini and Staniswalis, 1994). We extend this null hypothesis somewhat so that we can apply it to the data that motivates this paper. The alternative hypothesis, which includes unexpected measurement mechanisms, is not as simple as a partial linear model.

Comparison of replicated functional measurements involves hypothesis testing and smoothing. The goal in the analysis of these data is realistically portrayed by hypothesis testing

*Statistical Engineering Division, National Institute of Standards and Technology, 100 Bureau Drive, Stop 8980, Gaithersburg, MD 20899-8980, USA

Key words: Adaptive regression spline; Functional data analysis; Generalized linear model; Hypothesis testing; Partial linear model
because the causes of differences between replicates are thought to be known. Smoothing is needed because the shape of the function being measured is not known. The statistics literature reflects various aspects of combining testing and smoothing. First, smoothing that is adaptive over the domain of the independent variable (Cummins et al., 2001) is needed not only because the curvature of the function may vary with the independent variable but also because in our case, the counts observed vary by orders of magnitude. Second, applying less smoothing than knowledge of the shape of the function permits unnecessarily reduces the power of the test (Fan and Huang, 2001). Third, before a simple parametric form is assumed, its validity should be checked by exploratory analysis (Hall and Tajvidi, 2000). Fourth, adaptive smoothing that involves many data-based choices leads to some of the problems familiar in subset regression (Miller, 1990). Fifth, knowledge of mechanisms pertinent to the data, which is often available in physical sciences, can help with the inference. The data considered in this paper provide a context for bringing together all these aspects.

The data are measurements of a liquid containing particles. The measurements are of the particle-size intensity function, a function that when integrated over a size interval gives for that interval the mean particle count per unit volume of liquid. We gauge particle size by the radius of a sphere with volume equal to the particle volume. We denote this radius by \( r \) and denote the intensity as a function of radius by \( \lambda \). A measurement is itself a histogram connected to the size intensity through binning a realization of the corresponding inhomogeneous Poisson process. Replicate measurements differ not only in the Poisson-process realization but also by a scale-factor difference in the underlying intensities and a constant shift in the registration of the histogram boundaries. These known differences between replicate measurements describe the null hypothesis. Alternatives are that the underlying intensities differ in ways that unlike a scale factor are particle-size dependent. Replicate measurements with more than a scale-factor difference imply that the measurement process is affected by interesting, unknown mechanisms.

The approach discussed in this paper is based on representing the intensities by cubic splines with irregularly-spaced knots. Consider a set of knots that provides both an adequate and parsimonious cubic-spline representation of both intensities. Moreover, consider a subset of these knots that provides an adequate and parsimonious representation of the difference between the intensities. If the shift in histogram registration were given, then generalized linear modeling would lead to an hypothesis test (Firth, 1991). This test can be modified to account for shift estimation. This case of given knots, which is discussed in Section 3, provides a starting point for the case of knot selection.

Data-driven smoothing can be implemented by stepwise fitting of adaptive regression splines (Stone et al., 1997) as well as by other procedures. This paper does not contribute to the comparison of these procedures. Rather, in Section 4, we see where knot elimination (Kooperberg and Stone, 1992), a simple version of adaptive spline fitting, leads. We see that controlling false rejection of the null hypothesis in a procedure that involves choice among many sets of knots is a problem. In the literature on data-driven smoothing, the solutions offered are model selection by Akaike’s information criterion (AIC), Bayes information criterion (BIC), or cross validation. We have not implemented any of these solutions. However, we do note that for effectiveness in the current application, these solutions must be tuned to match the probabilistic aspects of the model and the knot comparisons that are performed at each step (Lindstrom, 1999; Zhou and Shen, 2001). Moreover, we note that data-driven smoothing with any of these model-selection criteria aims to select from a broad range of alternatives. For the application considered here, limitation of the alternatives through use
of prior knowledge would seem to have potential for being more effective.

In contrast to data-driven smoothing, which is a nonparametric approach, Section 5 offers a semiparametric approach based on prior knowledge available for the current application. The knowledge pertains to the plausibility of alternatives to the null hypothesis. Under both the null hypothesis and plausible alternative hypotheses, we model the difference between the two intensities parametrically. We model the function common to both intensities nonparametrically. Because the basis of the parametric part of the model is prior knowledge, checking on the parametric part with nonparametric modeling is, of course, appropriate. One would like an approach that blends nonparametric and semiparametric modeling. Put differently, exploring the nature of the difference should precede a formal test for it (Hall and Tajvidi, 2000). The current application shows that such blending may be possible on a case-by-case basis.

2. Particle-size application

The hypothesis-testing problem introduced in Section 1 is of genuine interest in the physical sciences. On the cutting edge of physical-sciences metrology is the measurement of functions. As in functional data analysis, functional measurements are observations that portray a function of a continuous independent variable. An example is measurement of a particle-size intensity. The challenge in functional measurement is replication, a cornerstone of science. The basic purpose of replication in metrology is discovery of unanticipated sources of variation. Such discovery can be framed as a hypothesis-testing problem in which the null hypothesis is that the difference between replicates can be attributed to the known sources of variation.

Functional measurements on Standard Reference Material 2806 (National Institute of Standards and Technology, 1997) illustrate the statistical concepts in this paper. The material consists of particles suspended in oil, more technically, ISO medium dust suspended in MIL-H-5606 hydraulic fluid. The dust, which is called Arizona road dust because of its origin, is often used for experiments in the fluid power industry. The measurement procedure begins with withdrawal of a 30 cm³ volume of the material from one of the bottles in which it is stored. This volume of material is then filtered leaving particles from this volume on a filter. Next, a scanning electron microscope (SEM) is used to obtain images of the filter surface at a random sample of locations. These images show the particles. Each particle that is evident is sized by counting the number of image pixels it covers. An image-wide threshold is the basis for determining whether a particle covers a pixel. Thus, each image produces a histogram of particle counts by particle area with area given as an integer number of pixels. The dust particles are sufficiently close to being spherical that the projected areas observed in an image are reasonable descriptors of size. We represent such an area by the radius of a circle of equal area. Finally, the histograms for the images are summed to give a histogram for the filter. For the purposes of this paper, the measurement procedure can be regarded as ending here although certification of the material for its intended purpose involves further steps.

Replicate measurements involve repetition of the entire measurement procedure starting with different bottles of the material. The known sources of variation are apparent in the procedure. The Poisson distribution seems reasonable because the filtering leaves the particles scattered on the filter according to what appears to be a mildly inhomogeneous Poisson process. The scale factor arises from the independent samplings of the different filters. The differences in registration of the histograms arises from differences in brightness
settings of the SEM. The settings are different for each filter but the same for the images of a filter. The brightness settings affect the number of pixels apparently covered by a particle and thus the histogram boundaries. Because this source of variation affects the edge of each particle, it is reasonable to suppose that the shift in histogram boundaries is a constant function of particle radius.

One can speculate over other potential sources of variation that hypothesis testing might expose. There are several places in the procedure where size fractionation is a possibility. The process of bottling the material, the process of extracting a volume from a bottle, and the process of filtering all might result in a measurement difference that cannot be characterized as simply a scale factor. The description of the effect of the SEM brightness setting leaves open the possibility that the registration of the histogram boundaries is not constant with radius. Finally, the measurement procedure is complex enough that it may not be fully understood.

Inspection of the data shows some aspects of note. First, only particles covering 6 or more pixels are counted in each image. Second, the particle-size intensity is roughly exponentially decreasing with radius. An important aspect of the 6-pixel cutoff is that its relation to actual particle size varies from replicate to replicate because of variation in SEM brightness. Third, the histogram bin boundaries are not uniformly spaced in radius. As a consequence of the intensity shape and the histogram spacing, the counts are higher for histogram bins on the left and less smoothing is needed on the left. Moreover, choice of a way to account for the histogram registration is different than if, for example, the histogram tailed off on the left.

3. Generalized linear models

The methodology of generalized linear models for the Poisson distribution with its canonical link function (Firth, 1991) is central to the application considered here. If an appropriate set of knots were given and if the histograms were identically registered, then a test for the difference between the intensities could be easily implemented with available software (Venables and Ripley, 1999). This section incorporates the shift in registration. What is known about the material measured is sufficient to allow specification of a set of knots large enough to represent the particle-size intensities. As is shown by the examples presented in this section, it may be possible to specify so many knots that the difference between the intensities cannot be detected.

A measurement consists of a sequence of counts, which we index by the corresponding numbers of pixels, 6, ..., and thereby tie our notation to a fundamental aspect of the data. We denote the counts for measurement $q$, $q = 1, 2$, by $y_{qj}$, where $j = 6, \ldots$. The radius for $y_{1j}$ given nominally in $\mu m$ is

$$r_j = \frac{1}{c}(j / \pi)^{1/2},$$

where $c = 0.5824$ for magnification 100× and $c = 2.856$ for magnification 500×. The radius scale is only nominally in $\mu m$ because there is no established relation between the threshold used to determine whether a pixel is a part of a particle and a physical definition of the particle edge. The radius for $y_{2j}$ is $r_j + \delta$, where $\delta$ is the shift in registration. For both measurements, the bin widths are

$$w_j = \frac{1}{\sqrt{2\pi c j}}.$$
Observations of a Particle-Size Distribution

The counts are assumed to be independently Poisson distributed. Denoting the intensities by \( \lambda_q(r) \), we have

\[
E(y_{1j}) = w_j \lambda_1(r_j)
\]

and

\[
E(y_{2j}) = w_j \lambda_2(r_j + \delta).
\]

Thus, the null hypothesis is that \( \lambda_1 \) is proportional to \( \lambda_2 \).

Some counts at the ends of the count sequences that make up the measurements do not enter the fitting of the generalized linear models. Moreover, treatment of the two measurements is not completely symmetric. There are three reasons for this. First, counts of particles covering just a few pixels may be uniquely affected by mechanisms that are of less interest because image processing failures for the smallest particles would not be surprising. Second, the number of particles observed becomes vanishingly small as the particle size increases. Third, inference on the shift \( \delta \) requires modeling with the same data set over the interval in which \( \delta \) might lie. For measurement 1, we base the analysis on counts for which \( r_L \leq r_j \leq r_U \), where \([r_L, r_U] = [r_{10}, r_{409}]\) for magnification 100\(\times\) and \([r_L, r_U] = [r_{11}, r_{609}]\) for magnification 500\(\times\). For measurement 2, we specify an interval for \( \delta_1 [\delta_L, \delta_U] \) and restrict the measurement 2 counts to those within the measurement 1 interval for all \( \delta \) in \( [\delta_L, \delta_U] \).

In other words, we select the measurement 2 counts in the set

\[
\{y_{2j} \mid r_L \leq r_j + \delta_L \text{ and } r_j + \delta_U \leq r_U\}.
\]

This is viable because \( \delta \) is small, some fraction of \( c^{-1}\mu \text{m} \), the length of the side of a pixel.

We introduce the cubic-spline representation of the intensities as follows: Let \( t \) be a \( k \)-dimensional vector containing the interior knots, and let \( t_0 \) be a \( k_d \)-dimensional vector containing a subset of the interior knots. Let \( (r - t)^3_+ = \max(0, (r - t)^3) \). For \( r_L \leq r \leq r_U \), we represent the sum and difference of the logarithms of the intensities as

\[
\log(\lambda_1) + \log(\lambda_2) = \sum_{i=0}^{3} a_i r^i + \sum_{i=1}^{k} a_{i+3} (r - t_i)^3_+
\]

and

\[
\log(\lambda_1) - \log(\lambda_2) = \sum_{i=0}^{3} a_{d(i)} r^i + \sum_{i=1}^{k_d} a_{d(i)+3} (r - t_{d(i)})^3_+.
\]

With this representation, the null hypothesis is that \( a_{d(i)} = 0 \) for \( i > 0 \).

For a given value of \( \delta \), the S-PLUS objects “bs”, “glm”, and “offset” can be used to fit either the null hypothesis or the alternative hypothesis. In the case of the null hypothesis, this fitting gives numerical values for all the model coefficients. In the case of the alternative model, if the interval \( [\delta_L, \delta_U] \) is too wide, some of the coefficients of the spline for the difference may not be determined. This indeterminacy can be corrected by removing knots from \( t_0 \).

Through fitting the model for various values of \( \delta \), we can estimate \( \delta \). Let \( y \) be a vector containing the counts from both measurements, let \( \lambda_0 \) be the corresponding intensity values under the null hypothesis, and let \( \lambda \) be the corresponding intensity values under the alternative hypothesis. For a given value of \( \delta \), S-PLUS gives the deviances corresponding
to the maximum-likelihood estimates for each model, $D(y; \hat{\lambda}_0, \delta)$ and $D(y; \hat{\lambda}, \delta)$. Minimizing $D(y; \hat{\lambda}_0, \delta)$ over $\delta$ in the interval $[\delta_L, \delta_U]$ gives $\hat{\delta}_0$, the maximum likelihood estimate under the null hypothesis. A similar approach under the more general alternative model is problematic. There will be little change in the deviance $D(y; \hat{\lambda}, \delta)$ with $\delta$ because the coefficients in the difference spline compensate for the change in $\delta$. Thus, when there are sufficient knots in the difference spline, assuming that $D(y; \hat{\lambda}, \delta)$ varies only a little with $\delta$ may be more reasonable than searching for $\hat{\delta}$ over a wide interval.

If a maximum likelihood estimate of $\delta$, $\hat{\delta}$, were available under the alternative hypothesis, then we could test the null hypothesis by referring the difference

$$D(y; \hat{\lambda}_0, \hat{\delta}_0) - D(y; \hat{\lambda}, \hat{\delta})$$

to the $\chi^2$ distribution with $k_d + 3$ degrees of freedom. Because of the confounding of the shift and the coefficients of the difference spline, we might substitute $\hat{\delta}_0$ for $\hat{\delta}$. If anything, this will make the statistic smaller and reduce the chance of rejecting the null hypothesis.

The four examples of function measurement comparison considered in this paper show different aspects of testing the null hypothesis. In the first example, there seems little one can do to justify rejection of the null hypothesis. In the fourth example, there seems no way to avoid rejecting the null hypothesis. In between, smoothing must be applied to reach a conclusion. For each example, we now present estimated intensities under the alternative hypothesis.

The result of fitting the alternative hypothesis to Example 1 is shown in Figure 1. Both log intensities are shown on top, and the difference is shown on the bottom. For this example, because $\hat{\delta}_0 = 0$, we do not offset one intensity with respect to the other. The knots included in $t$ and $t_d$ are shown. The knot spacing increases with radius as seems reasonable because the intensity decreases with radius. As one would expect from the decreasing counts, the estimated difference between the intensities fluctuates more with increasing radius. Taking into account the null hypothesis that one intensity is proportional to the other and thus that the log intensities differ by a constant amount, we see little in Figure 1 to suggest rejection of the null hypothesis. Moreover, the outcome of the generalized maximum likelihood test based on the difference of deviances gives a $p$ value of 0.782. Although as with all the examples, smoothing through estimation with fewer knots might reveal something more, applying such smoothing did not for this example. This example demonstrates that for the function measurements considered here, the null hypothesis must be taken seriously and not treated as just a statistically convenient reference point.

Example 2 is shown in Figure 2. This example differs somewhat from the others in that the two measurements are of the same filter. Thus, the difference does not involve the steps in extracting the particles from the hydraulic fluid and placing them on filters. This example is important because what is at issue here is whether the shift is constant with radius. The two measurements were made under different settings of the SEM brightness and modeling the effect of this is important. In the analysis that resulted in the certification of SRM 2806, modeling the shift as constant is key to combining data from different magnifications. On the basis that $\delta_0 = 0.39$, we have aligned the two intensities to eliminate the shift. Compared to Example 1, one can say that there is more visual evidence of size-dependent difference between the two intensities although the behavior of the difference for large radii is questionable because of the low counts. The generalized maximum likelihood test gives a $p$ value of 0.097. The next section details further investigation through additional smoothing.

Example 3 is shown in Figure 3. On the basis that $\hat{\delta}_0 = 0.117$, we have aligned the
Fig. 1: Log intensities (top) and log intensity difference (bottom) from initial smoothing of Example 1
Fig. 2: Log intensities (top) and log intensity difference (bottom) from initial smoothing of Example 2
Fig. 3: Log intensities (top) and log intensity difference (bottom) from initial smoothing of Example 3
two intensities. This example differs from the others in that the magnification is higher. Also, the number of knots is higher. In any case, visual evidence of difference between the intensities is not conclusive. The outcome of the generalized maximum likelihood test is a p value of 0.053. Further investigation through additional smoothing seems warranted.

Example 4 is shown in Figure 4. For this example, we take $r_U = r_{359}$ instead of $r_U = r_{409}$ because the difference in log intensities becomes so large for large particle sizes that the character of the difference for smaller particle sizes cannot be seen on the graph. On the basis that $\hat{\delta}_0 = -0.39$, we have aligned the two intensities. Here the generalized maximum likelihood test indicates rejection of the null hypothesis. Moreover, such rejection is indicated visually. In particular, the increase in the difference between the intensities in the vicinity of radius $11 \mu m$ is notable. It seems clear that a search for the cause of the evident lack of repeatability of the measurement process is justified.

4. Stepwise knot deletion

Further investigation of the difference between measurements calls for smoothing. Moreover, the wide range of observed counts calls for smoothing that is adaptive over particle size. This section illustrates adaptive smoothing through knot elimination implemented in a way that imitates LOGSPLINE (Kooperberg and Stone, 1992). Stepwise spline methods presented in the literature include an approach to selection of a set of knots from the sequence of knot sets generated through the progression of steps. We, however, do not formalize our approach to knot set selection. Rather, we just present the estimate for a knot set in the sequence.

In performing knot elimination, we begin with the general model with $t = t_d$ that is the basis for Figures 1-4 in Section 3. In each step, we eliminate one knot, the same one from $t$ and $t_d$. We eliminate the least significant knot as show by comparing fit of the model from the previous step with the fits of models with individual knots eliminated. We do knot elimination for $\delta$ fixed at $\hat{\delta}_0$ as obtained in the previous section. For examples 2 and 3, the ones presented in this section, at the step at which we stopped knot elimination, we re-estimated $t$ and found that it had not changed. This is not surprising in light of the shape of the log intensity as shown in Figures 2 and 3. For a log intensity with a different shape, a different strategy for dealing with the shift might have to be developed.

The result of smoothing the data in Example 2 is shown in Figure 5. With this set of knots, the p value for the test of the null hypothesis is 0.006. Looking at the supposedly significant difference between the two measurements in Figure 5, we see fluctuation about a constant level up to about 12 $\mu m$ followed by more pronounced variation. In light of what is known about the mechanisms that might cause replicate measurements to differ, one might judge that the fluctuations up to 12 $\mu m$ are the result of bin-to-bin Poisson variation alone. Yet, they contribute to the apparent significance of the difference. One possibility is that the smoothing process results in knots placed to emphasize contribution of the Poisson variation to the differences between the measurements. Thus, in stepwise smoothing coupled with a hypothesis test, one cannot just assume that the resulting knots were placed arbitrarily. Rather, one must recognize that the knot placement is data driven. In this sense, the smoothing coupled with hypothesis testing has aspects like those of variable selection in multiple regression (Miller, 1990).

The result of smoothing the data in Example 3 is shown in Figure 6. With the set of knots shown, the p value is 0.001. Once again, the difference shows considerable fluctuation. Some of this fluctuation is a contribution of the Poisson variation to the significance of the
Fig. 4: Log intensities (top) and log intensity difference (bottom) from initial smoothing of Example 4
Fig. 5: Log intensities (top) and log intensity difference (bottom) from data driven smoothing of Example 2
Fig. 6: Log intensities (top) and log intensity difference (bottom) from data driven smoothing of Example 3
test of the null hypothesis. One might easily look at Figure 6 and assert that the difference is under smoothed and that the method for selecting the knot set should be improved. But can the selection method be improved enough?

The question of model selection in a stepwise smoothing procedure is usually tied to the question of noise-related fluctuations in the resulting estimate. Here, the situation is analogous in that the noise-related fluctuations are in the difference. What is usually not considered in selection methods is a related hypothesis test. In the application considered here, what is needed is a knot set selection method that controls the rate of false rejection of the null hypothesis. Is it realistic to expect of find such a method without imposing prior knowledge of the alternative hypothesis?

Looking at Figures 5 and 6, one might wonder if one were to smooth further so that the seemingly spurious fluctuations were to disappear, whether the difference would still be significant. One basis for further smoothing is a general method for knot set selection such as AIC, BIC, and cross validation. Another basis for further smoothing is knowledge of what shapes of the difference are credible in terms of the subject matter origins of the data. Subject-matter knowledge provides an approach to multiple comparisons problems that amounts to focus on just some of the comparisons.

5. Semiparametric modeling

There is considerable understanding of the physical-sciences basis for the current application. In particular, causes for the difference between measurements can be listed with some assurance that the list is complete. For this reason, it is possible to judge the believability of a graph that portrays the difference. If there is a difference beyond what is accounted for under the null hypothesis, then it will likely be due to some mechanism that causes size fractionation or one that changes the images of the particles on the filter surface. The effect of such mechanisms can be expected to vary slowly with particle size. This is the reason why the fluctuations in the differences shown Figures 5 and 6 do not seem believable. The particle size intensity of the material itself might not be smooth but this does not show up in the difference.

An alternative to nonparametric smoothing is parametric modeling of the difference on the basis of physical understanding of plausible mechanisms. This leads to a combination of nonparametric modeling and parametric modeling, which is called semiparametric modeling. In the current application there are various ways that this might be done. One possibility would be to adopt a parameterization that forces the difference to be monotonic. This would be in keeping with the idea that size fractionation mechanisms cause monotonic differences. Of course, it is possible that two such mechanisms might have effects of roughly the same size and the result would not be monotonic. It should be noted that the proper parameterization for the difference might not be one that fits into a stepwise smoothing progression. That said, we note that for the examples, our approach is to assume that the difference is describable by a cubic spline with 0 or 1 interior knots, which does fit our stepwise progression.

Consider the case of 1 interior knot. In this case, the parameters are \( t_d \) (which is one dimensional), \( a_d(i), i = 0, \ldots, 4 \), and \( \delta \). Let \( a_d = (a_d(1), \ldots, a_d(4)) \). Consider the sum of the log intensities to be represented by a fixed set of knots given as components in of the vector \( t \), and let \( a \) be the coefficients of the cubic spline representation. Let \( \hat{a} \) be the maximum likelihood estimate of \( a \) for fixed \( t_d \), \( a_d \), and \( \delta \); and let \( D(y, \hat{a}, t_d, a_d, \delta) \) be the corresponding deviance. The estimate \( \hat{a} \) is a function of \( t_d, a_d, \) and \( \delta \). As a function the difference parameters, \( D(y, \hat{a}, t_d, a_d, \delta) \) is minus twice the log profile likelihood (up to
Observations of a Particle-Size Distribution

To estimate \( t_d, a_d, \) and \( \delta \), we minimize \( D(y; a, t_d, a_d, \delta) \). We compare the resulting minimized deviance with the minimized deviance under the null hypothesis to test the null hypothesis.

The fit for Example 2 under the alternative hypothesis is shown in Figure 7. This figure displays \( t \) by way of the locations of 19 interior knots, more than the number used in Section 4. The only knots not used here but used in Section 3 are ones that clearly do not contribute to the accuracy of the representation. In choosing \( t_d \), we minimized over the knots in \( t \). We could have moved the difference knot continuously but did not because moving it would not have made much difference. The shift under the alternative hypothesis is \( \hat{\delta} = 0.38 \), and the shift under the null hypothesis is \( \hat{\delta}_0 = 0.38 \). If we ignore the minimization over the position of the difference knot, the \( p \) value for the test of the null hypothesis is 0.071. We also note that in Figure 7, the shape of the difference is not completely credible at the right edge. Thus, we find it hard to reject the null hypothesis.

For Example 3, we drop the interior knot and its coefficient from our model of the alternative hypothesis. The fit for Example 3 under the alternative hypothesis is shown in Figure 8. This figure shows \( t \) by way of the locations of 33 interior knots. We obtain \( \hat{\delta} = 0.122 \) and \( \hat{\delta}_0 = 0.130 \). The \( p \) value for the test of the null hypothesis is 0.00009. Apparently, the changes in the difference between the two intensities from left to right in the Figure 8 reflect an unknown mechanism. An investigation of this mechanism seems justified.

Checking the prior knowledge applied in Examples 2 and 3 can be done by testing the fit of the parametric model for the alternative hypothesis. For Example 2, comparing the difference model with 1 interior knot with the difference model with the 19 interior knots that are shown in Figure 7, we obtain a \( p \) value of 0.066. If the parametric model is indeed backed by prior knowledge, then there seems to be no reason to reject it. For Example 3, comparing the difference model with no interior knots with difference model with the 33 interior knots that are shown in Figure 8, we obtain a \( p \) value of 0.387. This provides no evidence that the parametric model should be rejected.

What Examples 2 and 3 illustrate is that semiparametric modeling is valuable but that the prior knowledge on which it is based may not be as certain as one would like. Thus, for the application considered in this paper, a blending of semiparametric and nonparametric methods seems appropriate. Moreover, the conclusions reached about further investigation of unexpected mechanisms that affect the measurement process seem satisfactory. Thus, we conclude that on a case by case basis, blending of semiparametric and nonparametric methods may be possible.

Acknowledgements

R. A. Fletcher, J. R. Verkouteren, E. S. Windsor, D. S. Bright, E. B. Steel, and J. A. Small performed the experimental steps that gave the particle counts used for illustration in this paper.
Fig. 7: Log intensities (top) and log intensity difference (bottom) from semiparametric smoothing of Example 2
Fig. 8: Log intensities (top) and log intensity difference (bottom) from semiparametric smoothing of Example 3
REFERENCES


