Robust Estimation of Bioaffinity Assay Fluorescence Signals

Dimitris Glotsos, Jussi Tohka, Jori Soukka, Juhani T. Soini, Member, IEEE, and Ulla Ruotsalainen, Member, IEEE

Abstract—In this paper, the challenging problem of robust mean-signal estimation of a single-step microparticle bioaffinity assay is investigated. For this purpose, a density estimation-based robust algorithm (DER) was developed. The DER algorithm was comparatively evaluated with four other parameter estimation methods (mean value, median filtering, least square estimation, Welsch robust m-estimator). Two important questions were raised and investigated: 1) Which of the five methods can robustly estimate the mean bioaffinity signal? and 2) How many microparticles need to be measured in order to obtain an accurate estimate of the mean signal value? To answer the questions, bootstrap and coefficient of variation (CV) analyses were performed. In the CV analysis, the DER algorithm gave the best results: The CV ranged from 0.8% to 4.9% when the number of microparticles used for the mean signal estimation varied from 800 to 30. In the bootstrap analysis of the standard error, the DER algorithm had the smallest variance. As a conclusion, it can be underlined that: 1) of all methods tested, the DER algorithm gave the most consistent and reproducible results according to the bootstrap and CV analysis; 2) using the DER algorithm accurate estimates could be calculated based on 80–100 particles, corresponding to a typical assay measurement time of 1 min; and 3) the investigated bioaffinity signals contained a large number of outliers (observations that severely deviate from the majority of data) and therefore robust techniques were necessary for the mean signal estimation tasks.

Index Terms—Bioaffinity assays, robust clustering based on density estimation, robust estimation.

I. INTRODUCTION

A FLUORESCENCE bioaffinity assay is a technique for measuring the presence of a target molecule using a biochemical binding reaction. The target molecules are recognized and bound by specific fluorescent-labeled antibodies; by measuring the fluorescent emission of the labeling dyes, information can be obtained concerning the type and amount of the target molecules [1]. In ArcDia TPX assays, the complexes of target molecules and labeled antibodies are bound to microparticles, which are excited and measured one at a time [2]. Typically 80–120 particles can be detected during a 60-s measurement, the observation time for one microparticle being about 50 ms on average. The mean fluorescence signal value, which is proportional to the target molecule concentration, must be reliably estimated based on the limited number of observed microparticles. This is not straightforward, since the data may include outliers (observations that severely deviate from the majority of true data values) [3], [4]. The question then arises as to the best statistical method for robust bioaffinity signal estimation for data sets with small number of observations or substantial number of outliers. Robust estimation techniques [5] have been proven as powerful methods for reducing the variation in parameter estimation under such conditions.

Robust estimation, in a loose, nontechnical sense, is concerned with the fact that many assumptions commonly made in statistics (such as normality, linearity, independence) are at most approximations to reality [5]. In robust estimation, some assumptions about the input data are made but only the majority of the data and not all the data are presumed to behave according to these assumptions. These assumptions, e.g., about the parametric form of the probability distribution of the data, are often reminiscent of those of standard parameter estimation techniques. Standard parameter estimation techniques such as maximum likelihood or the least squares method, however, presume that all the data behave according to the assumptions in contrast to robust estimation techniques [5].

Traditional detection of outliers has been performed using variations of the David Hartley Pearson method that uses concepts of statistics to decide on the significance of each point to the final estimation [8]. Regression analysis and robust estimators have been widely studied [4], [6], [7] and used previously for outlier detection from normally distributed data [11] and multivariate data [12]. Regression analysis is a standard tool for parameter estimation in clinical chemistry that has been applied, for example, to immunoassays for prostate-specific antigen [13], for C-Reactive protein in blood samples [14], serum estradiol-17b [15], and mycophenolic acid [16]. Robust estimation techniques, on the other hand, have been shown to be superior to
regression analysis for data sets comprising small number of samples contaminated by outliers [6].

In this paper, we introduce a robust approach based on the concept of density estimation for robust parameter estimation of signals from a single-step microparticle-based bioaffinity assay. The goal was threefold: 1) to estimate robustly the fluorescent bioaffinity signal while isolating outlying measurements; 2) to investigate how small the particle population can be without significantly degrading the estimation performance; and 3) to compare the performance of the proposed density estimation algorithm with other methods from regression analysis and robust estimation that have been used in clinical chemistry applications [13]–[16].

II. MATERIAL AND METHODS

A. Bioaffinity Assay

The bioaffinity assay is based on ArcDia TPX-technology [2], [17], [18]. The TPX-technology uses the concept of two-photon excitation for detection of fluorescent polystyrene microspheres. The microspheres are used as the solid phase of the binding assay for capturing the analyte molecules and the fluorescent antibody molecules. In addition to the detection antibody fluorophore, each microsphere possesses internal coding fluorophores in the longer wavelength emission region for identification of microsphere subclasses. An assay may contain several microsphere subclasses for detection of multiple analytes simultaneously. A tightly focused laser beam defines a microvolume for two-photon excitation. Due to the high beam intensity, optical forces are generated that pull microspheres to the center of the excitation zone. The passage of microspheres through the focal region causes a significant change in the backscattered light, which is detected using a confocally located photodetector. This scattering signal activates the fluorescence measurement. The generated fluorescence emission is recorded by three photomultipliers, each one adjusted to measure the emission from a specific labeling dye. A more detailed explanation of the assay technology can be found in [2], [17], and [18].

The scattering and fluorescence signals measured during the transit of each microsphere through the focal region encode significant information concerning the type, kinetics, and concentration of the binding analyte [17]. However, due to variations of the fluorescence labeling and the complexity of the measurement system, the fluorescence signals are contaminated with outliers. Accurate signal estimation greatly depends on eliminating outliers, thus, robust estimation of these signals is of high importance.

For each microparticle, three parameters were utilized for the robust estimation of the mean bioaffinity fluorescence: 1) the bioaffinity fluorescence signal; 2) the scattering signal; and 3) the observation time of the microsphere.

B. Density Estimation-Based Robust (DER) Algorithm

The establishment of concrete a priori assumptions concerning the underlying distribution of bioaffinity signals is very difficult, if not impossible, since one measurement obtained from a microsphere may differ from another due to instrumentation complexity, originating mainly from variation in microsphere searching and optical trapping.

Considering this, we designed and applied an algorithm recently published by our group [19], [20] to approximate the average bioaffinity signal based on the nonparametric estimation of its underlying distribution. The DER algorithm starts by estimating the data probability density function (PDF) \( f : \mathbb{R}^d \to \mathbb{R} \) at each data point \( x_i, i = 1, \ldots, N \) using a Gaussian Parzen window [21], [22] according to

\[
f(x_i) = \frac{1}{N(2\pi)^{d/2}\sigma^d} \sum_{j=1}^N \exp \left[ -\frac{(x_i - x_j)^2}{2\sigma^2} \right]
\]

where \( N \) is the number of data points, \( d \) is their dimensionality, and the parameter \( \sigma \) defines the spread of the applied Gaussian window function. Although the optimal value of \( \sigma \) is usually experimentally determined, in this paper we estimated \( \sigma \) according to

\[
\sigma = \frac{1}{(N-1)^2} \sum_{i=1}^N \sum_{j=1}^N \|x_i - x_j\|, \quad i \neq j.
\]

In this way \( \sigma \) is independent of the type of data and additionally \( \sigma \) is automatically estimated. Following this, the whole data PDF is estimated based on the individual PDF estimates \( f_i = f(x_i) \). The most prominent peak of the distribution is defined using a maximum likelihood estimation-based algorithm [23]. This peak is considered as the data centroid. The algorithm continues by assuming that the most distant points of the PDF (see point \( O_{\text{right}} \) and \( O_{\text{left}} \) in Fig. 1) are the most likely outliers. Data points in the PDF that are closer to \( O_{\text{right}} \) or \( O_{\text{left}} \) than the data centroid are then considered outliers, by means of a \( k \)-nearest neighbor heuristic. Outliers are then rejected and the final signal estimation is performed by scaling the remaining data points according to their distance from the data centroid. Identifying two extreme points as outliers, even if they are not outliers, leads only to a small loss in the statistical accuracy.
of the mean signal estimation. This loss can be quantified by comparing the variance of the estimates obtained with the sample size of $N$ and the sample size of $N-2$, the difference of which is usually insignificant. A more detailed description of the DER algorithm and its implementation in a neural network configuration can be found in [19], and [20].

C. Other Mean Signal Estimation Methods

The average bioaffinity signal was additionally estimated using four other parameter estimation methods. For each microparticle, input data consisted of the bioaffinity fluorescence signal and the observation time that were measured during the transit of each microparticle through the excitation zone. Signal estimation methods were the following.

1) The mean value of the bioaffinity signal was calculated as the arithmetic mean after filtering out the measurements whose observation time was outside the range 5–120 ms.

2) The mean value of the bioaffinity signal was calculated after filtering out measurements whose value was outside the range of twice the standard deviation from the median and the measurements whose observation time was outside the range 5–120 ms.

3) The relationship between the fluorescence signal and the observation time was quantified by performing a least squares (LS) analysis. The average bioaffinity signal was calculated as the mean value of the LS fit [24] after filtering out measurements outside the range 5–120 ms.

4) Using the Welsch robust $m$-estimator, the robust fit between the bioaffinity signal and the observation time was calculated. The average signal was then computed considering as outliers those points outside the range of two times the standard deviation from the median.

In the following paragraphs, we will describe the robust Welsch estimator in more detail, since it represents modern robust estimation techniques, and hence it is the method that the DER algorithm should be compared to.

The classical LS method is a well-established approach for parameter estimation [9] and has been traditionally used for signal analysis of bioaffinity assays [13]–[16]. However, the LS estimators are based on certain assumptions, such as that random errors follow a specific distribution, are uncorrelated, have the same variance, and have zero mean. These assumptions are rarely met in experimental science. Particularly, the squared error criterion of the LS estimate is sensitive to the existence of outliers, since the objective function error criterion of the LS estimate is sensitive to the existence of outliers. Particularly, the squared error criterion of the LS estimate is sensitive to the existence of outliers, since the objective function error criterion of the LS estimate is sensitive to the existence of outliers.

The average signal was additionally estimated using some other criterion instead of the least square error. The effect of outliers is reduced by replacing the squared error criterion with a more slowly increasing function of residuals. There is no “golden” criterion for selecting the best $\rho$-function for general purposes. The selection is application and data specific. A “rule of thumb” is to select the estimator based on the data distribution [5]. In this study, we not only evaluated the Welsch estimator, but also tested other $M$-estimators including Andrews, bisquare, Cauchy and, Huber estimators [5]. The $\rho$-function for the Welsch estimator is given by

$$\rho_{\text{Welsch}}(x) = \frac{c^2}{2} \left[ 1 - \exp \left( -\left( \frac{x}{c} \right)^2 \right) \right]$$

where $c$ is a fixed tuning parameter equal to 2.9846. To solve (3) an iterated re-weighted least squares (IRLS) algorithm was used [25]. The iteration is only guaranteed to converge to a unique solution for convex $\rho$-functions and redescending functions, but in the case of the Welsch $\rho$-function, (3) may have multiple roots. This means that the starting point of the iteration must be chosen carefully. The initialization in this study was obtained by using the Huber minimax non-descending function, that has been proven [26] to globally converge to a unique solution. Additionally, there is a scaling problem with the Welsch estimator, since it depends on a scale factor. To rescale data, the Median of absolute deviation (MAD) estimator of the scale was used [5].

D. Bootstrap Analysis

To test and compare the functionality between the DER algorithm and the remaining four estimation algorithms, we performed bootstrap and coefficient of variation (CV) analyses. In part, our aim was to find out what is the minimum number of particles that needs to be measured for the accurate estimation of the mean bioaffinity signal. According to the bootstrap method [27], we randomly drew subsets of $k$ particle signals, computed the average fluorescence signal, repeated the same procedure $n$ times, and calculated the coefficients of variation, standard errors, and the confidence intervals of the estimators. This was done using several $k$ values (microparticle measurement populations of 800, 700, 600, 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, and 30, respectively) tested for four different $n$ values (50, 100, 1000, and 10 000) and for all robust methods to get a reliable picture of each algorithm’s functionality. The CV is defined as the standard deviation (of the estimates) divided by the mean (of the estimates).
III. RESULTS AND DISCUSSION

A. Demonstration of the Data Nonnormality

In Fig. 2, the normal probability plot of the data (for all 874 microparticle measurements) is displayed. Sampled data are illustrated with “+” whereas the line (calculated as a robust linear fit of the sample order statistics) is used to evaluate the linearity of the data. If data follow a normal distribution, the plot will appear linear. However, this is not the case, since the plot clearly indicates the nonnormality of the underlying distribution, suggesting that there are outliers among the data. The rejection of these outliers is crucial for reliable parameter estimation of these bioaffinity signals, making robust techniques vital for this application.

The nonnormality of the data distribution was additionally verified by performing the Jarque–Bera test [28]. According to this test, the hypothesis of data having a normal distribution with unspecified mean and variance is evaluated, starting with the null hypothesis that data does not have a normal distribution. The test estimates the data skewness and kurtosis, which should be near 0 and 3, respectively, in the case of a normal distribution. The p-value of the test was 2.7%, the value of the test statistic was 334.4, and the critical value for rejecting the null hypothesis was 1 indicating that data follow a nonnormal distribution.

B. DER Example

The performance of the DER algorithm in identifying possibly outlying measurements for the bioaffinity signal with respect to observation time is demonstrated in Fig. 3. This figure illustrates which data in the “distribution space” were translated as outliers in the “signal-to-time-in-focus space” using the DER algorithm. The results were obtained by analyzing measurements of 874 microparticles without bootstrap. The input to the DER algorithm comprised the bioaffinity and the scattering signals.

The main advantage of the proposed DER algorithm compared to robust estimation methods that extend the idea of LS estimation, is that its setting is much easier when more than one observation is investigated [29]. Additionally, the algorithm needs no special formulation for multidimensional input, and its computational burden is relatively moderate, since no iterative procedures or feedbacks are required. For the specific application, the DER algorithm was implemented in custom software developed using MATLAB. The algorithm required less than 5 s for detecting outliers and calculating the mean bioaffinity signal in the case of a one-dimensional input signal (bioaffinity fluorescence) and approximately 70 s for two-dimensional input (bioaffinity fluorescence and scattering signal).

C. Results of the Bootstrap Analysis

Results after performing CV analysis using the bootstrap method for different microparticle populations \( k \) and for \( n = 10000 \) are presented in Fig. 4. Additionally, standard errors and confidence intervals were computed (see Fig. 5 and Table I). CV analysis enables us to determine the tradeoff between result precision and total assay measurement time. As it can be observed in Fig. 4, the best precision is achieved using about 800 microparticles (CV less than 1%), which corresponds to approximately 8 min of total assay measurement time. Typical particle detection rate is 80–120 particles per min.

By reducing the number of microparticles, the particle-to-particle variation increases exponentially. Results with the DER algorithm seem consistently better than with any other method for measurements down to 100 microparticles. The latter has to be stressed, considering that the particle-to-particle variation for \( k = 100 \) with the DER method is relatively small, down to 2.29%. In this way, by ensuring consistent results with only 100 microparticles measurements, total assay measurement time can be reduced to approximately 1 min.
The bootstrap standard error was calculated as the standard deviation of the mean value estimate for each bootstrap replication (Table I) [27]. The DER algorithm presented the smallest error for all k microparticle populations examined. Ninety percent, 95%, and 99% confidence intervals were calculated to give an estimate of the range of values that is likely to include the average bioaffinity signal computed with each method. Confidence intervals were calculated using the percentile method, i.e., by estimating the mean value from k-samples with an estimator and then computing the 2.5 and 97.5 percentiles. Results for the width (subtraction of the upper from the lower limit) for the 95% confidence interval are demonstrated in Fig. 5 for the Welsch and DER algorithms. The width of the confidence interval gives us some idea about how uncertain we are about the unknown parameter (average bioaffinity signal). The DER algorithm presents almost the same width with the Welsch estimator for the 95% confidence interval from 800 to 100 microparticles, but has the narrowest interval for k from 100 to 70 possibly indicating that it can calculate the mean bioaffinity signal value with more certainty for smaller particle populations. All the confidence intervals for the remaining three methods were significantly wider. The dramatic increase of the width for populations smaller that 70 microparticles (80 for the Welsch estimator) indicates that the estimation results are inaccurate in over 5% of cases. This can be explained by the fact that the number of outliers, as a ratio of the data, increases as k-shrinks after a critical value of k. This is the reason why we selected to perform bootstrap and CV analysis: to study the behavior of the estimation methods in small populations and to get an idea how many microparticles need to be measured in order to obtain an accurate estimate of the mean signal value. The distributions for the different bootstrap data sets and for different particle populations are illustrated in Fig. 6, where it is clear that the discrepancy between data sets grows when k shrinks.

The bootstrap simulation was additionally performed for lower values of n (50, 100, and 1000) and the outcome was similar, indicating that the DER algorithm was the most consistent of the estimation methods studied. Moreover, the results with the other M-estimators tested (Andrews, bisquare, Cauchy, and Huber) were inferior to the results with the Welch estimator that are listed here.

According to the preceding bootstrap and CV analysis the DER algorithm gave the most consistent and reproducible results compared to the other methods tested. Calculation is relatively fast (approximately 2 min), enabling the extraction of consistent results with microparticle population measurements down to 100. Hence, total assay measurement time can be reduced to approximately 1 min, while ensuring the consistency and reproducibility of the estimation.
IV. CONCLUSION

In this paper, we assessed the performance of five different algorithms on the task of robustly estimating fluorescent signals collected from a single-step microparticle-based bioaffinity assay. Two important questions were raised and investigated in this research: 1) Which of the five methods can robustly estimate the mean bioaffinity signal? and 2) How small the microparticle population can be without significantly degrading the estimation performance? The answers to these questions are not straightforward. However, the proposed DER algorithm presented the most attractive and robust statistical results after performing bootstrap analysis. As a conclusion, it can be underlined that: 1) the investigated bioaffinity signals contained a large number of outliers (observations that severely deviate from the majority of data) and therefore robust techniques are necessary for the mean signal estimation tasks; 2) of all methods tested, the DER algorithm gave the most consistent and reproducible results according to the bootstrap analysis; 3) setting and complexity of the DER algorithm is much easier in the case of more than one-dimensional observations compared to the methods tested; and 4) results indicated that the DER algorithm could support stable performance (for $k = 100$ and $n = 10000$, CV was 2.29%) with even 80–100 microparticle population measurements, which is translated to total assay measurement time down to 1 min. The latter is most important since in immunodiagnostic applications, acceptable CV cannot exceed 5% and tolerated measurement time ranges from 20 to 60 s.

REFERENCES

Dimitris Glotsos received the B.Sc. degree in medical instruments technology from the Technological Institution of Athens, Athens, Greece, in 2000, and the M.Sc. and Ph.D. degrees in medical physics from the University of Patras, Patras, Greece, in 2002 and 2006, respectively. Currently, he is a Marie Curie Fellow Researcher at the Institute for Genomics and Bioinformatics, Graz University of Technology, Graz, Austria.

Jussi Tohka received the M.Sc. degree in pure mathematics from the University of Turku, Turku, Finland, in 1999 and the Ph.D. degree (with honors) in signal processing from Tampere University of Technology, Tampere, Finland, in 2003. He visited the Montreal Neurological Institute (fall 2001–spring 2002) and was a Postdoctoral Fellow at the Laboratory of Neuro Imaging, University of California, Los Angeles. Currently, he is a Postdoctoral Fellow at the Institute of Signal Processing, Tampere University of Technology. His research interests include statistical pattern recognition, image analysis, and applications of them in biosciences, particularly in brain imaging.

Jori Soukka received the M.Sc. degree in electronics from the University of Turku, Turku, Finland. Currently, he is with Arctic Diagnostics Oy, Turku, as a Senior Engineer.

Juhani T. Soini (M’02) received the M.Sc. degree in biomedical engineering from Tampere University of Technology, Tampere, Finland, in 1997, and the Ph.D. degree in medical physics from the University of Turku, Turku, Finland, in 2002. He has been working in several analytical instrumentation and biotech companies. Currently, he is working as the Professor of medical physics and engineering at the University of Turku, and is the Head of the Finnish DNA Microarray Center, University of Turku and Åbo Akademi University, and holds the position of Adjunct Professor of biomedical instrumentation at the University of Turku.

Ulla Ruotsalainen (M’98) received the M.Sc. degree in electronics engineering in 1976 and the Ph.D. degree in signal processing in 1998 from Tampere University of Technology, Tampere, Finland. She has also received the degree of Medical Physicist in 1990. Currently, she is a Professor of signal processing at the Institute of Signal Processing, Tampere University of Technology.