KALMAN SMOOTHING OF GLUCOSE DATA

APPLIED TO PARTIAL LEAST SQUARES MODELING OF NON-INVASIVE NEAR-INFRARED MEASUREMENTS

Odd Martin Staal[1,2,3,4], Steinar Sælid[4], Terje Karstang[4], and Øyvind Stavdahl[1,2]

odd.martin.staal@prediktor.no steinars@prediktor.no terje@prediktor.no oyvind.stavdahl@itk.ntnu.no

[1]Department of Engineering Cybernetics (www.ntnu.no/itk)
[2]Artificial Pancreas Trondheim – The APT research group (www.apt-norway.com)
[3]The Research Council of Norway (www.forskningsradet.no/)
[4]Prediktor Medical AS, Fredrikstad, Norway (www.prediktormedical.com)

Norwegian University of Science and Technology (NTNU), Trondheim, Norway

MOTIVATION

- Glucose data from Self Monitoring Blood Glucose Meters (SMBG) , Continuous Glucose Monitors (CGM) and laboratory equipment provide heterogeneous data sets, where differences in data formats, measurement error and sampling frequency can compromise efficient and correct use of the data
- A useful preprocessing step is converting the data into continuous signal of blood glucose (BG) values and uncertainties.
- The variance estimate can indicate which data sets are too sparse for a particular application, or allow automatic removal of intervals where the information is not sufficient.
- Interpolation of glucose data is useful in Partial Least Squares (PLS) modelling of Near Infrared (NIR) data, where it can be used to generate a response value for each NIR sample . A commonly used alternative to the method presented here is to use cubic spline interpolation, which may introduce some artefacts.

METHODS

Kalman smoothing of glucose data

Discrete-time glucose data is processed by a Kalman smoother to produce a continuous series of estimates (mean and variance).

Glucose (G) and its rate ($\dot{G}$) is modelled as $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$ with $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & -a \end{bmatrix}$ and $\boldsymbol{x} = [G \quad \dot{G}]^{\mathrm{T}}$

Discretization yields $\boldsymbol{x_{k+1}} = \boldsymbol{\Phi x_k}$ with $\boldsymbol{\Phi} = \boldsymbol{e^{A\Delta t}}$. We used Δt = 0.5 minute. A value of a = -0.05 was found to give satisfactory interpolation in the test data.

The measurement variance R depends on what kind of glucose measurements y are available. For YSI laboratory measurements, R = 10^{-4} y^2 $(mmol/L)^2$ is appropriate and according to the precision listed for the instrument.

For SMBG meters, R can be set to 0.172 for y <= 5.6 mmol/L, and to $0.01y^2$ for y > 5.6 mmol/L, where y is the measured value. These are worst case variances for SMBG devices fulfilling [1]. The variances found in [2] are somewhat lower.

$Q = \begin{bmatrix} 0 & 0 \\ 0 & 0.005 \end{bmatrix}$ was used as the process covariance matrix and was found by tuning to give realistic variances in the estimates.

The glucose data are Kalman filtered using the glucose measurements at the points in time where they are available, and at all other times the filter only performs time updates. A backward pass is performed according to the Rauch-Tung-Striebel formulas [3], using the estimates and the covariance matrices from the forward pass.

Response variable for NIR measurements

The algorithm was used to generate a response variable for PLS modelling of NIR data collected with the BioMKR® device (pictured below). This device generates 10 NIR responses every 30 seconds, from a measurement system consisting of 10 NIR LEDs and a photodiode.

The NIR and glucose data investigated were collected in a clinical trial where venous blood was sampled every 5 minutes by a YSI 2300 instrument during a hyper- and hypoglycemic clamp.

Other data sets containing SMBG measurements collected during clinical and volunteer testing of the BioMKR® device have also been used in testing the method.

RESULTS

The figures on the right show the estimates from the Kalman smoother (blue) with variance estimates (blue shadow, 2σ). A cubic spline interpolation (red) of the same data is shown for comparison. Several applications of the method are illustrated.

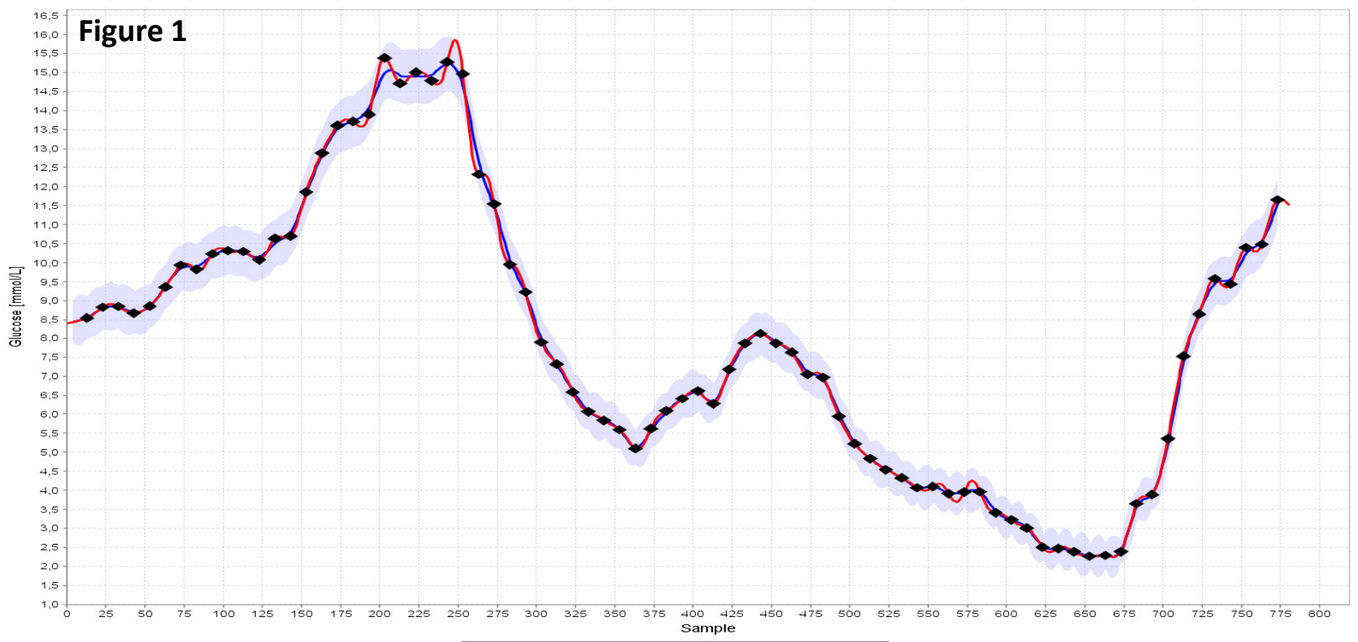
Figure 1 shows automatic correction of readings from a clinical trial clamping study with glucose measurements from a YSI Stat 2300 instrument. The oscillations introduced by cubic spline interpolation are eliminated.

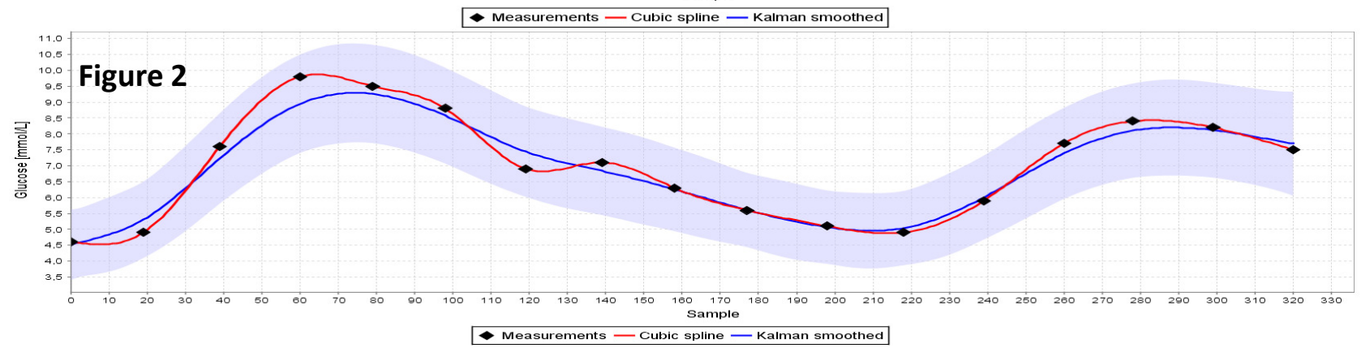
Figure 2 shows more sparsely sampled SMBG data with larger measurement error, resulting in a larger estimated variance in the smoothed output.

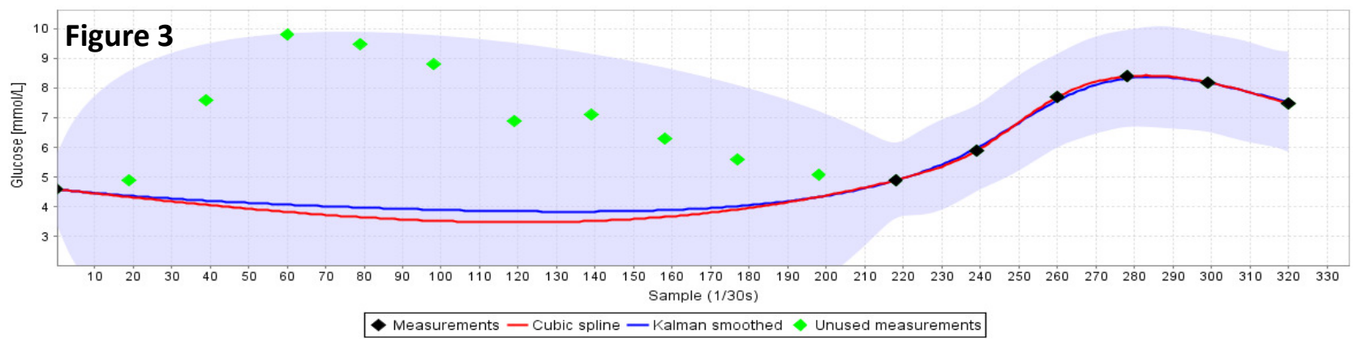
Figure 3 shows the same data set as that of Figure 2, but with a simulated data loss in the first part of the series; the unused measurements are shown in green. This illustrates the realism of the variance of the estimate.

When the Kalman smoother is used to generate the response variable for PLS modeling of the BioMKR NIR signals, an improvement in Standard Error of Calibration (SEC) and Standard Error of Validation (SEV) is seen in almost all data sets when compared to using cubic spline interpolation for the response generation, the average SEV dropping from 1.46 (cubic spline) to 1.39 (Kalman smoothed) in the 24 NIR data sets tested. Two possible reasons are hypothesized:

1. Cubic spline interpolation introduces synthetic and irrelevant fluctuations that may affect PLS calibration adversely.
2. The smoothed data better approximates the real response seen by the NIR measurement, which in theory should be a mixture of blood and interstitial glucose, thus it would not fluctuate as much and as rapidly as blood glucose does.







CONCLUSION

Kalman smoothing of glucose data is useful to convert discrete time glucose data sets into continuous data with a measure of uncertainty (variance) at each point in time, and has inbuilt robustness against measurement errors and missing data.

When used to generate a response curve for PLS modeling of NIR glucose data, the Kalman smoothing gives a better fit to the data than cubic spline interpolation.

REFERENCES

[1] ISO 15197:2013 In vitro diagnostic test systems -- Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus

[2] Vettoretti, M.; Facchinetti, A.; Sparacino, G. & Cobelli, C. Accuracy of devices for self-monitoring of blood glucose: A stochastic error model *Conf Proc IEEE Eng Med Biol Soc,* **2015***, 2015*, 2359-2362

[3] Fredrik Gustafsson, 2010, Statistical Sensor Fusion, Studentlitteratur AB, Lund