Akademie věd České republiky
Ústav teorie informace a automatizace, v.v.i.

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

Václav Šmídl

Robust Detection of Linear Part of Patlak-Rutland Plots

No. 2243 December 2, 2008

ÚTIA AV ČR, P.O.Box 18, 182 08 Prague, Czech Republic
Tel: (+420)266052422, Fax: (+420)286890378, Url: http://www.utia.cas.cz,
E-mail: utia@utia.cas.cz

1 Introduction

Patlak-Rutland (PR) plot is a mathematical tool used in analysis of dynamic nuclear medicine studies [2]. This graph plots a curve which is supposed to contain a linear part. The length and position of this linear part is unknown. The main advantage of PR plot is visualization of the curve which allows manual selection of the linear part by an expert.

In this report, we propose an automated method of selection of the linear part using probabilistic modelling and Bayesian statistic. In order to do that we review the model of renal activity that is observed on the dynamic nuclear medicine studies.

1.1 Mathematical model of renal function

We study function of a kidney using the following curves that are observed on scintigraphic studies:

$Q(t)$ activity in renal (kidney) region of interest with subtracted background,

$P(t)$ activity in blood,

where t are discrete times of measurement.

The following model is considered to hold approximately:

$$Q(t) = k \sum_{\tau=1}^{t} P(\tau) + hP(t) + \sigma(t)e(t). \tag{1}$$

where k and h are unknown constants, and $e(t)$ is an accumulated error of approximation.

Model (1) is in principle a linear regression model with two notable exceptions:

- the model holds only for $t \in [t_l, t_h]$ where t_l and t_h are *unknown* lower and upper bounds of the interval. Lower and upper bounds on those unknowns, $\bar{t}_l < t_l < t_h < \bar{t}_h$, are known from physical model:

$$\begin{aligned} \bar{t}_l &= \min(\max_t P(t),\ 1\text{min}). \\ \bar{t}_h &= \min(\max_t R(t),\ 3\text{min}). \end{aligned} \tag{2}$$

- $e(t)$ can be considered as zero-mean Gaussian distributed, however temporal evolution of variance is uknown. We will consider two hypothesis:

  constant variance: $\sigma(t) = \sigma$ in equation (1).

  Patlak-Rutland: variance of $e(t)$ is proportional to $P(t)^{-2}$, $\frac{\sigma(t)}{P(t)} = \sigma$.

Naming of the second hypothesis is motivated by the RP plot,

$$\frac{Q(t)}{P(t)} = k\frac{\int P(t)dt}{P(t)} + h + \frac{\sigma(t)}{P(t)}e(t), \tag{3}$$

which arise from (1) by dividing both sides by $P(t)$.

The task is to find all unknown variables: t_l, t_h, k, h.

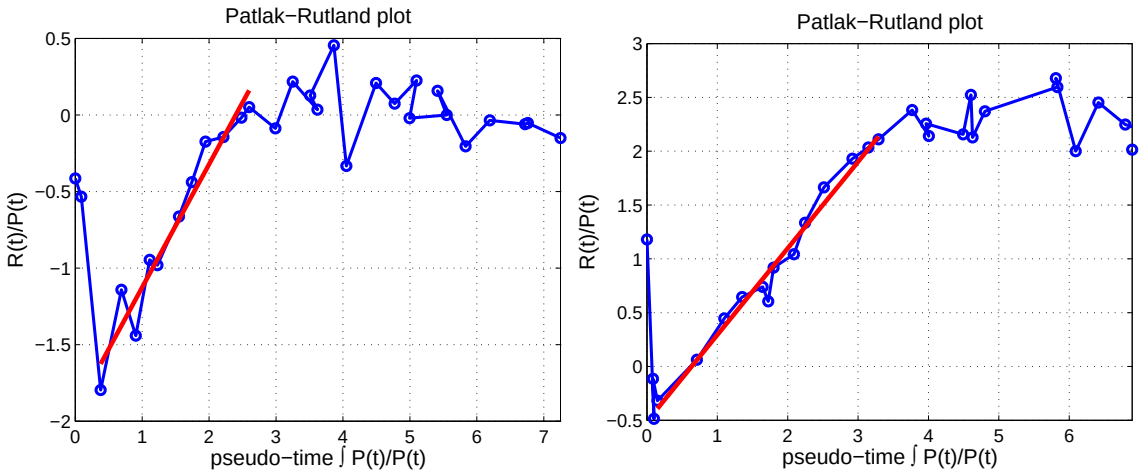

Figure 1: Two examples of Patlak-Rutland plots with fitted linear segment.

2 Bayesian Estimation of Linear Regression Parameters

Under the assumption of Gaussian noise, models (1) and (3) can be rewritten in the form

$$f(y_t|\psi_t, \theta, \sigma) = \mathcal{N}(\theta\psi_t, \sigma^2), \tag{4}$$

using assignments $\{y_t = Q(t), \psi_t = [\sum P(t), P(t)], \theta = [k, h]\}$ for model (1) and $\{y_t = Q(t)/P(t), \psi_t = [\frac{1}{P(t)} \sum P(t), 1], \theta = [k, h]\}$ for model (3).

Using Bayesian framework, posterior density of the model parameters is found using the Bayes rule:

$$f(\theta, \sigma|y_{t_l} \dots y_{t_h}, \psi_{t_l} \dots \psi_{t_h}) = \frac{f(\theta, \sigma) \prod_{t=t_l}^{t_h} f(y_t|\psi_t, \theta, \sigma)}{f(y_{t_l} \dots y_{t_h}, \psi_{t_l} \dots \psi_{t_h})},$$
$$f(y_{t_l} \dots y_{t_h}, \psi_{t_l} \dots \psi_{t_h}) = \int f(\theta, \sigma) \prod_{t=t_l}^{t_h} f(y_t|\psi_t, \theta, \sigma) d\theta d\sigma.$$

where $f(\theta, \sigma)$ is the prior density.

Model (4) belongs to the exponential family of models, for which sufficient statistics exist if the prior is chosen form the conjugate family. Then, posterior density of the model parameters is found in the form

$$\begin{aligned} f(\theta, \sigma|y_{t_l} \dots y_{t_h}, \psi_{t_l} \dots \psi_{t_h}) &= \frac{f(\theta, \sigma) \prod_{t=t_l}^{t_h} f(y_t|\psi_t, \theta, \sigma)}{f(y_{t_l} \dots y_{t_h}, \psi_{t_l} \dots \psi_{t_h})}, \\ &= GiW(V, \nu). \end{aligned} \tag{5}$$

where

$$\begin{aligned} V &= V_0 + \sum_{t=t_l}^{t_h} \begin{bmatrix} y_t \\ \psi_t \end{bmatrix} [y_t \, \psi_t'], \\ \nu &= \nu_0 + (t_h - t_l + 1). \end{aligned} \tag{6}$$

Statistics V_0, ν_0 are statistics of the prior denstity which is of the same form as posterior (5). Decomposing V into four parts

$$V = \begin{bmatrix} V_y & V_{y\psi} \\ V_{\psi y} & V_\psi \end{bmatrix}, \tag{7}$$

the following important results can be derived [3]:

$$\hat{\theta} = V_{\psi y} V_\psi^{-1} \tag{8}$$

$$f(y_{t_l} \dots y_{t_h}, \psi_{t_l} \dots \psi_{t_h}) = \frac{I(V,\nu)}{I(V_0,\nu_0)} \tag{9}$$

$$I(V,\nu) = (2^{\nu+2}\pi^{2-t_h+t_l})^{\frac{1}{2}}\Gamma(\frac{\nu}{2})|V_\psi|^{-\frac{1}{2}}|V_y - V_{y\psi}V_\psi^{-1}V_{\psi y}|^{-\frac{\nu}{2}}. \tag{10}$$

Here, α_0 is normalization integral of the prior.

3 Fitting Linear Interval

Since identification of the linear regression parameters for known t_l and t_h is trivial, we adress identification of the time bounds first. We seek probability $f(t_l, t_h|\mathcal{M}_i, Q, P)$ where $\mathcal{M}_i$ denotes conditioning on knowledge of the model of the data (1) or (3) for $i = 1, 2$, respectively. Using the chain rule of probability

$$\begin{aligned} f(t_l, t_h|\mathcal{M}_i, Q, P) &= \int f(t_l, t_h, k, h, \sigma|\mathcal{M}_i, Q, P)dk\, dh\, d\sigma \\ &\propto \int f(Q(t_l \dots t_h), k, h, \sigma|\mathcal{M}_i, P)dk\, dh\, d\sigma \times f(Q_{\text{out}}) = \\ &= f(Q(t_l \dots t_h)|\mathcal{M}_i, P) f(Q(t_{\text{out}})|P). \end{aligned} \tag{11}$$

Here, $t_{\text{out}} = [\bar{t}_l \dots t_l - 1, t_h + 1 \dots \bar{t}_h]$ denotes data outside interval $[t_l, t_h]$. In order to proceed, we need to complement model $\mathcal{M}_i$ by models of the data outside of the linear region. In this paper we consider uniform distribution of the data

$$f(Q_{\text{out}}) = \prod_{t \in t_{\text{out}}} U(q_{\min}, q_{\max}) = \left(\frac{q_{\min} - q_{\max}}{K_q}\right)^{|t_{\text{out}}|}, \tag{12}$$

where $q_{\min}$ and $q_{\max}$ are apriori known lower and upper bounds of the admissible range, K_q denotes proportion of the range in which the points can reside, see Fig 2 for illustration.

Typical number of records in the considered region is around 20. Thus extensive testing of all possible combinations of t_l and t_h involves around 200 evaluations of (11) which is computationally cheap.

Using (9) in (11) and removing all elements independent of t_l, t_h, V and ν, the posterior is

$$\begin{aligned} f(t_i, t_j|\mathcal{M}_i, Q, P) \propto &(2^{\nu+2}\pi^{2-t_h+t_l})^{\frac{1}{2}}\Gamma(\frac{\nu}{2})|V_\psi|^{-\frac{1}{2}}|V_y - V_{y\psi}V_\psi^{-1}V_{\psi y}|^{-\frac{\nu}{2}} \\ &\left(\frac{q_{\min} - q_{\max}}{K_q}\right)^{-|t_{\text{out}}|}. \end{aligned}$$

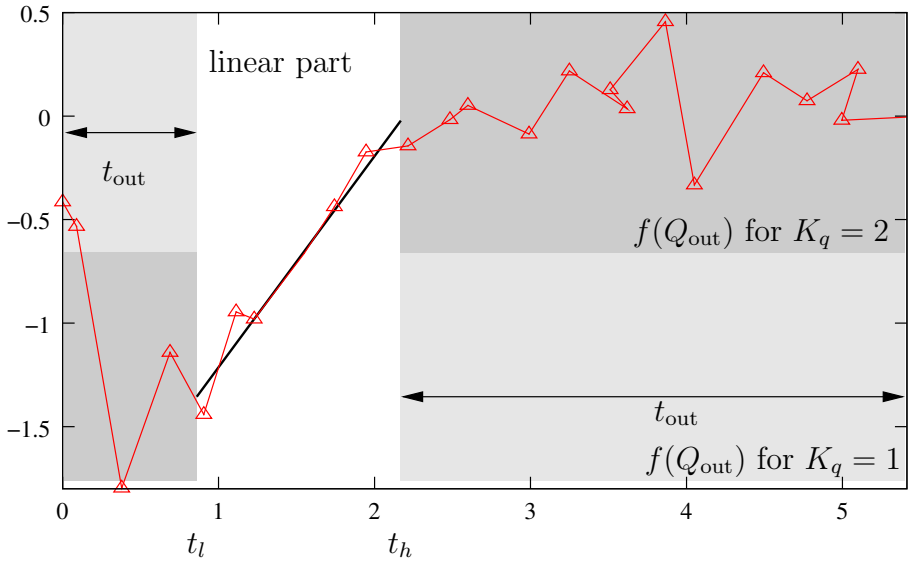

Figure 2: Illustration of model of the points outside the linear interval. Grey area covers the uniform model where external points can be anywhere between $q_{\min}$ and $q_{\max}$, i.e. model (12) for $K_q = 1$. Dark grey area covers the uniform model where external points can be only in one half of the interval $q_{\min}$ and $q_{\max}$, in this case chosen in lower part of the first and upper part of the second external region.

3.1 Model combination

It remains to address, which of the two models—(1) and (3)—is more appropriate. A simple solution arise when we assume both models to be equaly probable. Then,

$$f(t_i, t_j|Q, P) \propto 0.5 f(t_i, t_j|\mathcal{M}_1, Q, P) + 0.5 f(t_i, t_j|\mathcal{M}_2, Q, P), \tag{13}$$

which can be easily computed.

3.2 Experimental results

The methods were tested on 20 data sets of real data for $K_q = 1$. The best estimates of unknown t_l and t_h are displayed in the left part of Table 1, and the corresponding estimates of the $\hat{\theta}$ are displayed in the right part of the table.

We note the following:

- The considered models yield identical estimates of the bounds t_l, t_h for 14 of 20 data files. For the remaining 6 data files, the proposed combination (13) favored 3 times bounds obtained for (1), while those for (2) were preferred only 2 times. An interesting case is the file `mag3_4i1c.crv` for which the combination picked different bounds from both competing models. The results are shown in detail in Figure 3.

- Estimates of the regression coefficients θ obtained for models (1) and (2) differ only negligibly when the bounds are estimated equally. Otherwise, they may differ substantially.

- For 6 data files, estimated bounds of the linear interval are at the fixed bounds for both considered models.

file name	$\bar{t}_l$	$\bar{t}_h$	eq. (1)		eq. (2)		eq. (13)		$\hat{\theta}$ for (1)	$\hat{\theta}$ for (2)
dtpa_1i1c.crv	2	16	2	0	2	0	2	0	0.56 -0.19	0.57 -0.20
dtpa_2i1c.crv	2	18	3	0	2	0	3	0	0.27 -0.80	0.25 -0.74
dtpa_3i1c.crv	2	18	3	0	3	0	3	0	0.34 -0.36	0.34 -0.35
dtpa_4i1c.crv	2	18	4	0	4	0	4	0	0.09 -0.04	0.09 -0.05
dtpa_5i1c.crv	2	15	2	0	2	0	2	0	0.46 -0.18	0.47 -0.19
mag3_1i9c.crv	3	18	1	0	1	0	1	0	0.87 -0.94	0.88 -0.96
mag3_2i1c.crv	2	18	0	0	0	0	0	0	2.18 -0.43	2.19 -0.45
mag3_3i1c.crv	2	14	2	0	0	0	0	0	1.06 -0.18	1.05 -0.15
mag3_4i1c.crv	2	15	2	0	1	-2	4	-2	0.64 -1.68	0.81 -1.94
r0001.crv	2	15	0	0	0	0	0	0	1.07 0.09	1.06 0.11
r0002.crv	3	18	0	0	0	0	0	0	0.44 -1.20	0.46 -1.25
r0003.crv	3	18	0	0	0	0	0	0	1.30 -0.29	1.32 -0.33
r0004.crv	2	11	2	0	2	0	2	0	1.24 -0.47	1.25 -0.49
r0005.crv	2	17	0	0	0	-1	0	-1	0.49 -0.26	0.50 -0.27
r0006.crv	2	18	5	0	1	0	5	0	0.91 -1.23	0.82 -0.97
r0007.crv	2	18	3	0	3	0	3	0	0.33 -0.95	0.33 -0.95
r0008.crv	4	16	0	0	0	0	0	0	0.77 -0.45	0.80 -0.51
r0009.crv	2	18	0	0	0	0	0	0	1.11 -0.24	1.12 -0.25
r0010.crv	2	18	2	0	0	0	2	0	0.61 -0.69	0.56 -0.54
r0011.crv	2	18	1	0	1	0	1	0	0.84 -0.58	0.84 -0.58

Table 1: Table of indexes of the best linear fits for all data files. Fixed upper and lower bounds, $\bar{t}_l$ and $\bar{t}_h$, are displayed in the first column. Differences of these bounds and the best fits of linear model for eq. (1) are in the second, while those for eq. (3) (Patlak-Rutland plot) are in the third. The proposed combination of those two approaches, eq. (13), is displayed in the third column. Estimates of the regression parameters corresponding to the used model are displayed in the last column.

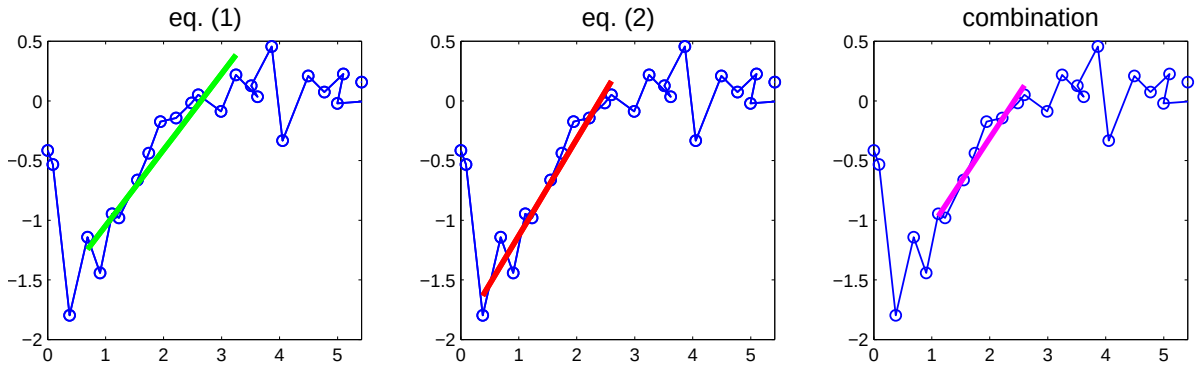

Figure 3: Estimates of the best linear fit using three different approaches for data file `mag3_4i1c.crv`.

- The method is very sensitive to the choise of K_q. For example, for a realistic choise $K_q = 3$, in 6 cases out of 20, the best linear part estimated by the method is only 3 points long.

4 Outlier suppression

Note that PR plots may contain outliers, Fig. 1 right. A way to suppress these outliers is to relax the assumption of a compact interval of linear fit. Specifically, we define two sets of time indexes t_{in} and $t_{\text{out}} = \{t : t \notin t_{\text{in}}\}$ such that

$$f(Q, P|\mathcal{M}_i) = f(Q(t_{\text{in}})|\mathcal{M}_i, P) f(Q(t_{\text{out}})|P). \tag{14}$$

Here, we again consider model (12) for observation in t_{out}. Extensive evaluation of all possible combinations of membership of time indexes in these sets is computationally prohibitive. Hence, we adopt an approximate evaluation procedure of the EM-like type. Consider a label process

$$l(t) = \begin{cases} 1 & t \in t_{\text{in}}, \\ 0 & t \in t_{\text{out}}. \end{cases}$$

Model (14) can be rewritten as follows:

$$f(Q, P|\mathcal{M}_i, l(t), \theta, \sigma) = \prod_{\forall t} f(Q(t)|\mathcal{M}_i, P, \theta, \sigma)^{l(t)} f(Q(t)|P)^{1-l(t)}.$$

If $l(t)$ were known, posterior density of θ, σ is again in the form (5) with statistics

$$\begin{aligned} V &= V_0 + \sum_{t=\bar{t}_l}^{\bar{t}_h} l(t) \begin{bmatrix} y_t \\ \psi_t \end{bmatrix} [y_t \, \psi_t'], \\ \nu &= \nu_0 + \sum_{t=\bar{t}_l}^{\bar{t}_h} l(t). \end{aligned} \tag{15}$$

Marginal density is then given by (9).

Using the idea of the EM algoritm [4], we can iteratively replace $l(t)$ by maximum likelihood estimate

$$\hat{l}^{(i)} = \arg\max_{l} \mathsf{E}_{\theta,\sigma|Q,P,l^{(i-1)}} \ln f(\theta, \sigma|Q, P, \mathcal{M}_i, l_i), \tag{16}$$

using posterior density from the previous iteration.

Since evaluation of (16) is non-trivial we approximate it as follows. We replace $l(t)$ by its expected value using estimate of $\hat{l}_{qb}^{(i-1)}$ for the remaining values of t, i.e.:

$$\begin{aligned} \hat{l}_{qb}^{(i)}(t) &= \mathsf{E}(t \in t_{\text{in}}|l(\tau) = \hat{l}_{qb}^{(i-1)}(\tau), \forall \tau \neq t, Q, P) \\ &= \frac{I(V_{t\in t_{\text{in}}}, \nu_{t\in t_{\text{in}}})}{I(V_{t\in t_{\text{in}}}, \nu_{t\in t_{\text{in}}}) + I(V_{t\notin t_{\text{in}}}, \nu_{t\notin t_{\text{in}}})}. \end{aligned}$$

where $V_{t\in t_{\text{in}}}$ denote V computed with $l(t) = 1$ and $V_{t\notin t_{\text{in}}}$ with $l(t) = 0$ respectively. This approach is closely related to the Quasi-Bayes approach of used in estimation of paremetes of probabilistic mixtures [1].

Note that:

file name	$\hat{\theta}$ for (1)	$\hat{\theta}$ for (2)	relative difference
dtpa_1i1c.crv	0.560 -0.178	0.564 -0.189	0.8% 5.8%
dtpa_2i1c.crv	0.271 -0.830	0.264 -0.809	2.6% 2.5%
dtpa_3i1c.crv	0.295 -0.208	0.294 -0.208	0.4% 0.4%
dtpa_4i1c.crv	0.064 0.008	0.067 0.002	4.1% 116.1%
dtpa_5i1c.crv	0.484 -0.211	0.488 -0.217	0.9% 2.8%
mag3_1i9c.crv	0.874 -0.929	0.887 -0.963	1.6% 3.6%
mag3_2i1c.crv	2.198 -0.456	2.204 -0.466	0.3% 2.3%
mag3_3i1c.crv	1.021 -0.103	1.013 -0.086	0.8% 17.8%
mag3_4i1c.crv	0.810 -1.962	0.775 -1.885	4.4% 4.0%
r0001.crv	1.064 0.120	1.067 0.115	0.3% 3.9%
r0002.crv	0.501 -1.313	0.508 -1.320	1.4% 0.5%
r0003.crv	1.289 -0.291	1.294 -0.303	0.4% 3.7%
r0004.crv	1.270 -0.502	1.291 -0.539	1.6% 7.0%
r0005.crv	0.503 -0.278	0.499 -0.269	0.9% 3.3%
r0006.crv	0.925 -1.276	0.917 -1.256	0.8% 1.6%
r0007.crv	0.334 -0.967	0.330 -0.959	1.1% 0.9%
r0008.crv	0.821 -0.517	0.824 -0.527	0.5% 2.0%
r0009.crv	1.118 -0.254	1.118 -0.257	0.0% 1.0%
r0010.crv	0.571 -0.587	0.569 -0.580	0.4% 1.2%
r0011.crv	0.854 -0.599	0.862 -0.623	0.9% 3.9%

Table 2: Estimates of the linear regression coefficients for models (1) and (2) using assumption (14).

- For this model, we do not have to set hard limits on the external points. They will be automatically suppressed if they do not fit the linear interval inside. However, this also means that even valid points may be rejected if the algorithm is initialized poorly – e.g. when the first and the last points are initialized with 1 and the rest with 0.

- Information that the points close to the limits (2) are most likely outliers is transformed into prior density on l_t as follows:

$$f_0(l(t)) = 0.99 \left(\sin(\frac{t}{T}\pi) + 0.1 \right).$$

  Thus, the first and the last point within the interval (2) have prior probability of belonging to the linear part approx. 0.1. While the point in the middle of the interval has prior probability of being in the linear part around 0.99.

4.1 Experimental results

The same 20 data files were tested using this model. Visualization of the results in the form of PR plot is displayed in Figure 4.

The most notable result is that for this approach the estimated parameters for model (1) and (2) differ only negligibly which is demonstrated by olmost perfect overlap of green and red lines in Figure 4 and numerical results in Table . Relative difference of these two estimates is

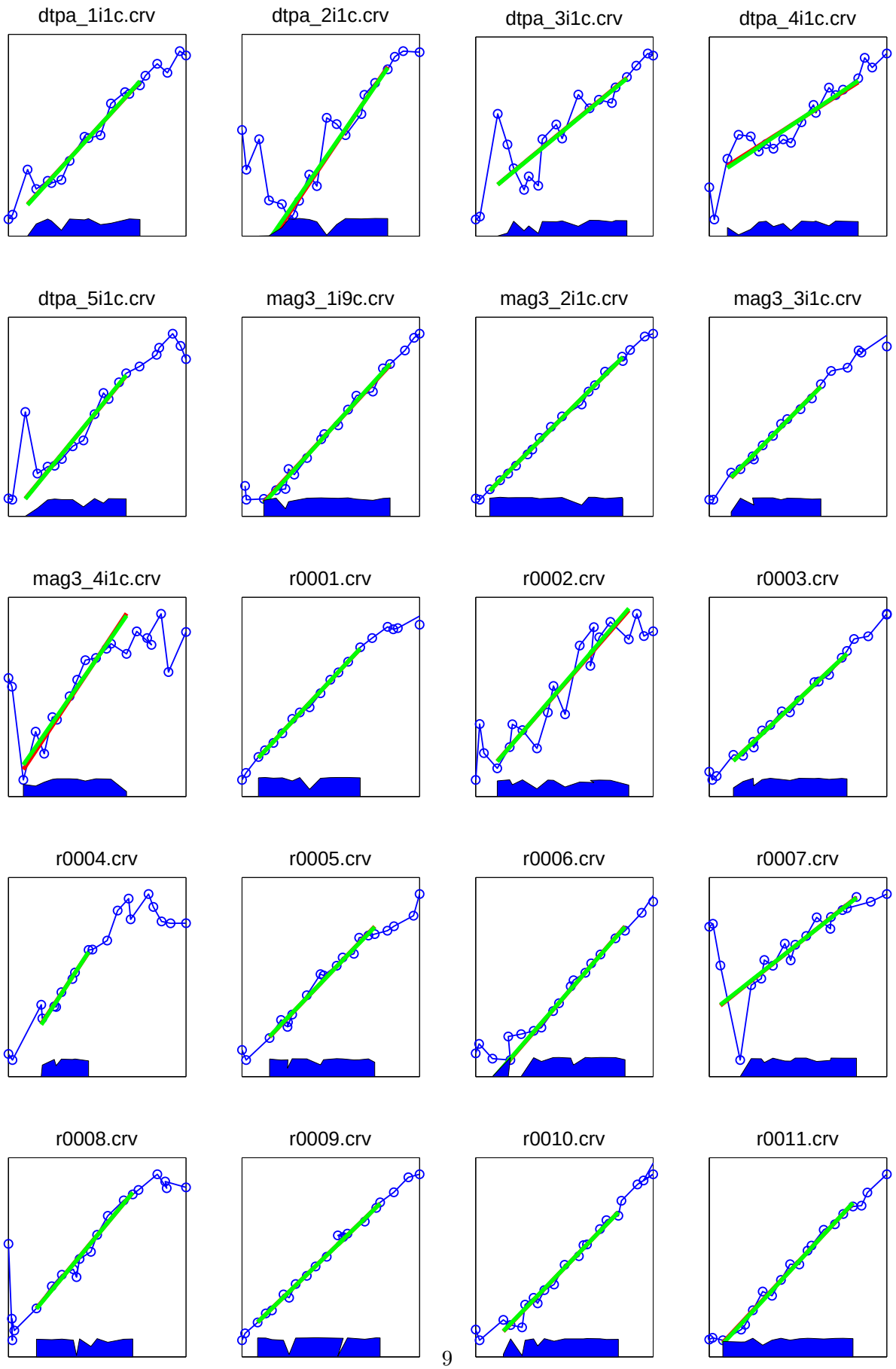

Figure 4: Linear fits of a linear model with outliers for 20 data sets. The best fit using model (1) is displayed in red, the best fit using model (2) is in green (however they mostly overlap). The thick blue bar st the bottom of each plot is proportional to the weight of the corresponding data point on the x-axis. Dips in this bar signify suppressed data, i.e. the outliers.

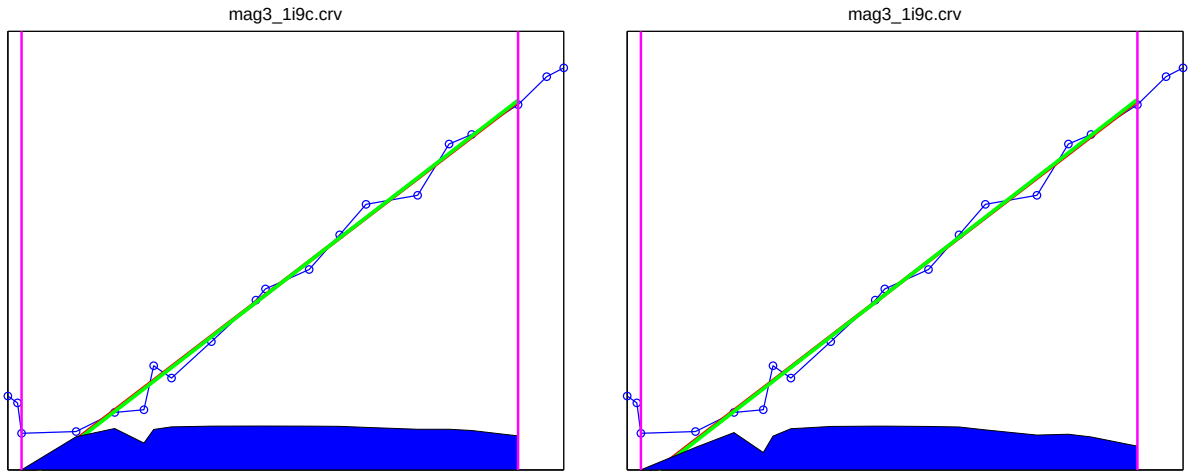

Figure 5: Linear fits for different values of K_q: $K_q = 1$ (left) and $K_q = 3$ (right).

typically lower than 7%. The only two exceptions from this rule are associated with numbers clos to zero, hence even if the relative error is high, the absolute error is minor.

This method is much less sensitive to the choise of K_q. Higher values of K_q results in more aggresive outlier suppression, Fig. 5. However, the resulting linear fits are almost identical. For example, coefficients of linear fits for the two plots in Fig 5 are: $0.877, -0.943$ and $0.874, -0.937$ respectively.

Matlab code for these experiments can be downloaded from: http://staff.utia.cas.cz/smidl.

5 Conclusion

Estimation of linear coefficients for functional analysis of renal activity form scintigraphic images is complicated by two facts: (i) it is not known which of the two competing models is more appropriate in given context, and (ii) the assumption of linearity is valid only within time-interval with unknown bounds. We have proposed two solutions based on Bayesian estimation theory. The first approach evaluates likelihood of all admissible bounds of the time-interval, the second approach seeks the best linear fit on the whole admissible data robust to outliers.

Both of the proposed methods eliminate complications mentioned above. Experimental comparison of the approaches on real data verified that the resulting estimates are robust to the choice of the competing models, yielding similar estimates for both of them. For this limited study, the approach based on outlier suppression appears to be more robust to the choise of the model and prior parameter K_q. Thus, it may be a good candidate for automatic or semi-automatic analysis of this problem. For example, one estimate for each model can be computed and when the procedure would be completed automatically if they are close to each other. When they differ significantly the case would be brought to the attention of an expert.

Acknowledgements

Support of grant MŠMT 1M0572 is gratefully acknowledged.

References

[1] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař. *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer, London, 2005.

[2] R.S. Lawson. Application of mathematical methods in dynamic nuclear medicine studies. *PHYSICS IN MEDICINE AND BIOLOGY*, 44:57–98, 1999.

[3] V. Peterka. Bayesian approach to system identification. In P. Eykhoff, editor, *Trends and Progress in System identification*, pages 239–304. Pergamon Press, Oxford, 1981.

[4] D.M. Titterington, A.F.M. Smith, and U.E. Makov. *Statistical Analysis of Finite Mixtures*. John Wiley, New York, 1985.