The Pinch Force Data

February 5, 2001

Conditions when collecting the data

Each subject was

- to maintain to within reasonable limits a constant force determined by the experimenter,
- upon a signal to give a brief force impulse which was targeted to reach a pre-determined maximum value, and
- to return the force to the constant background level.

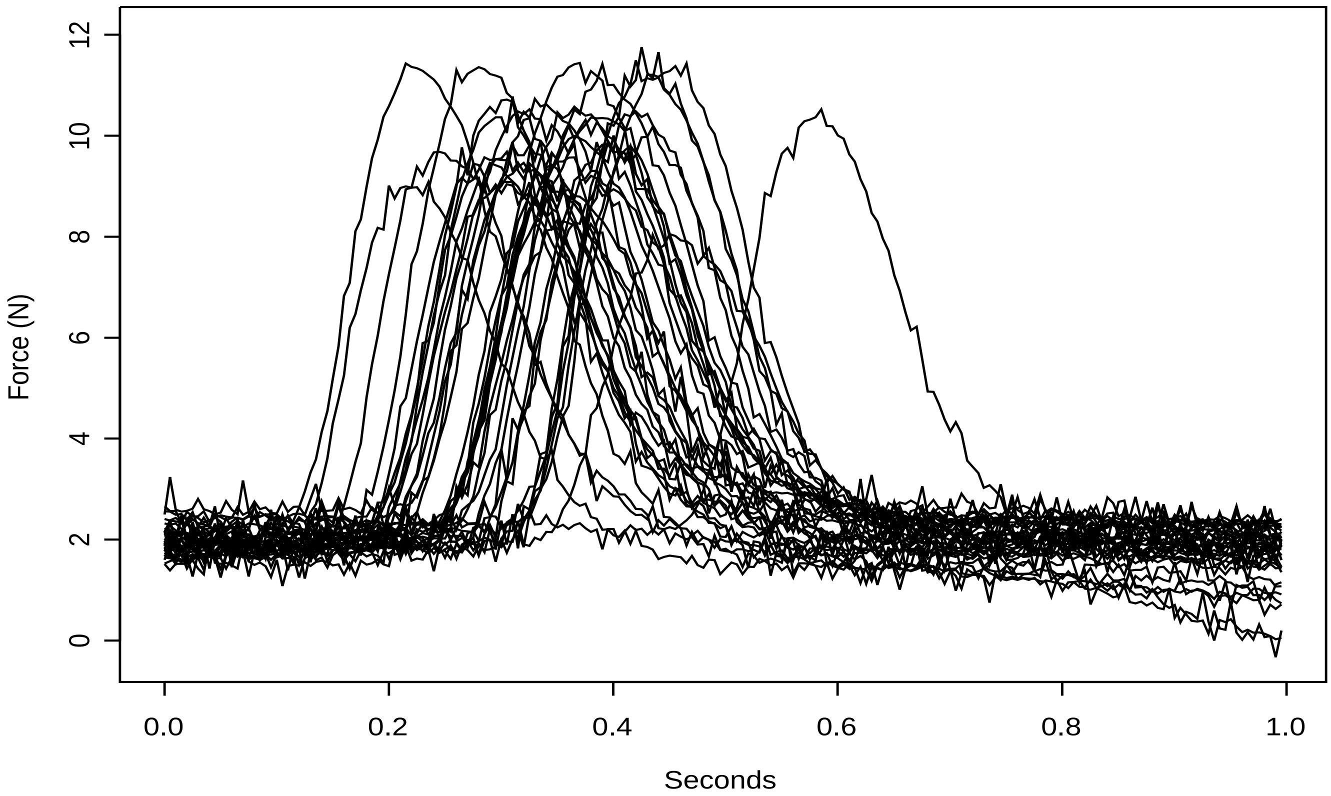


Figure 1: Original 36 records

Observation:

- this is an example of functional data, i.e. the basic or primitive observation is a function, the force as function of time.

Objective:

- find a representation of this process that determine the typical record as well as the variation between records, like a simple function or a simple differential equation, and so on.

Method 1:

- apply principal component analysis directly on the raw data set.

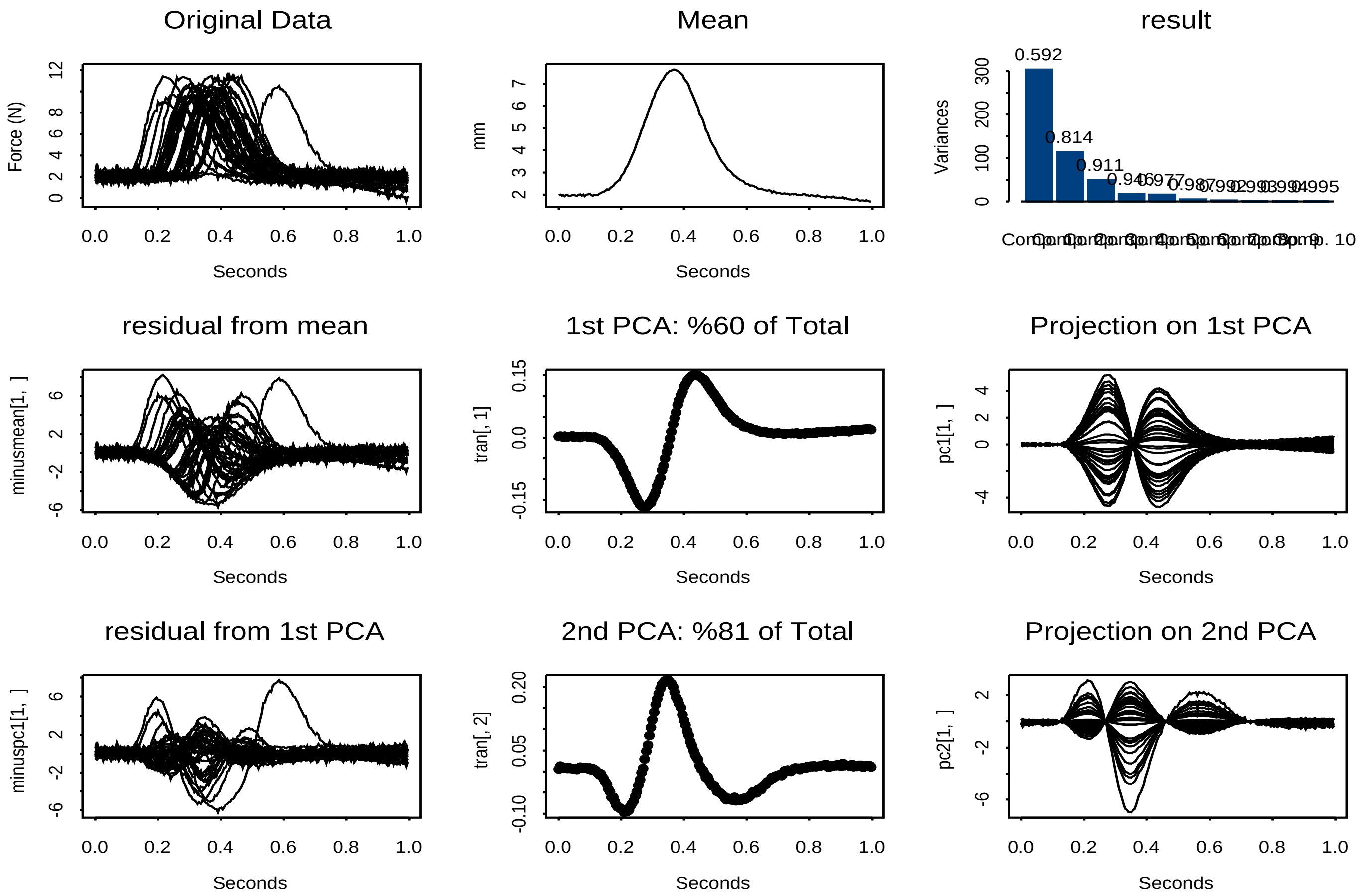


Figure 2: Fist picture of PCA

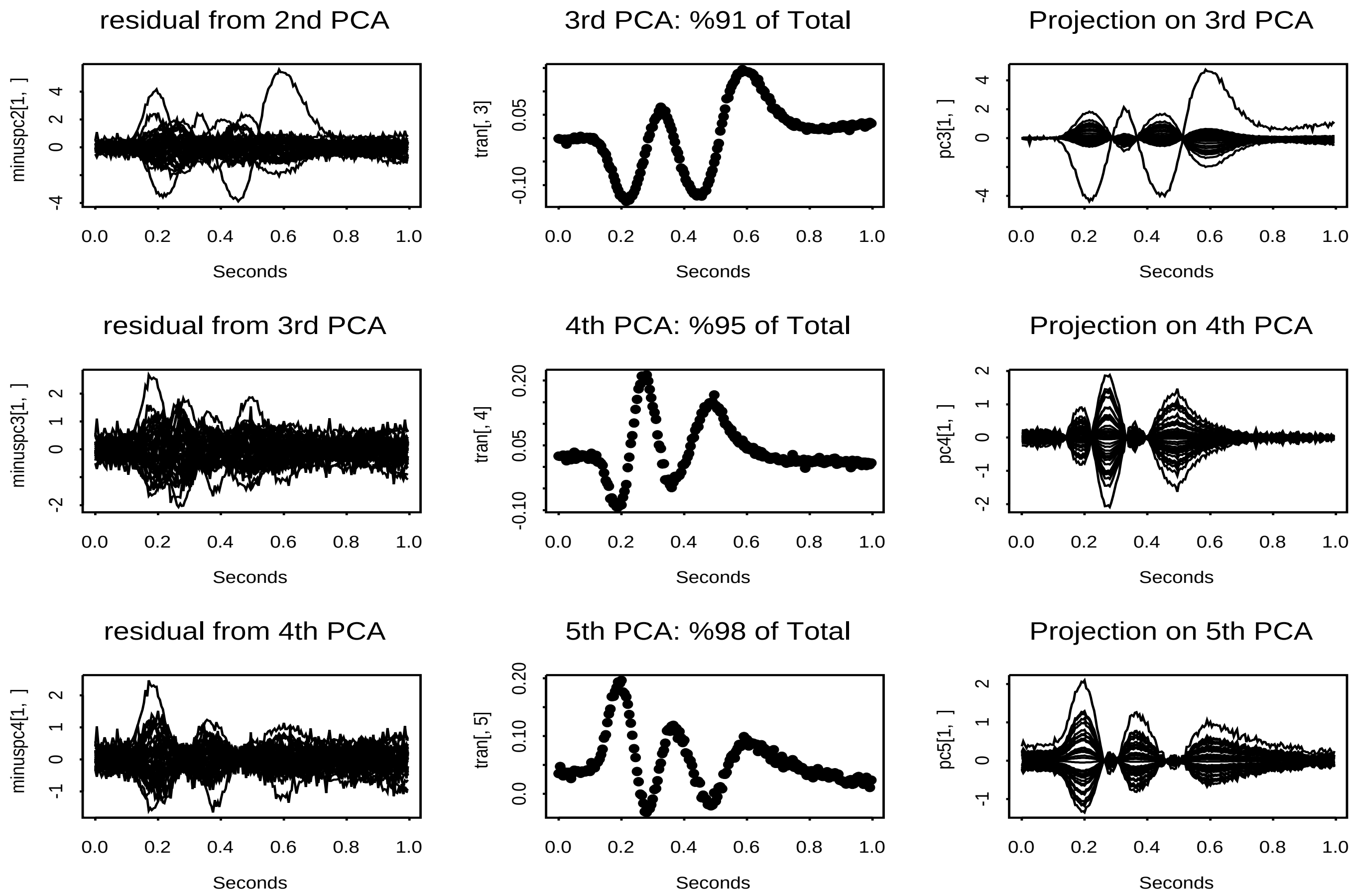


Figure 3: Second picture of PCA

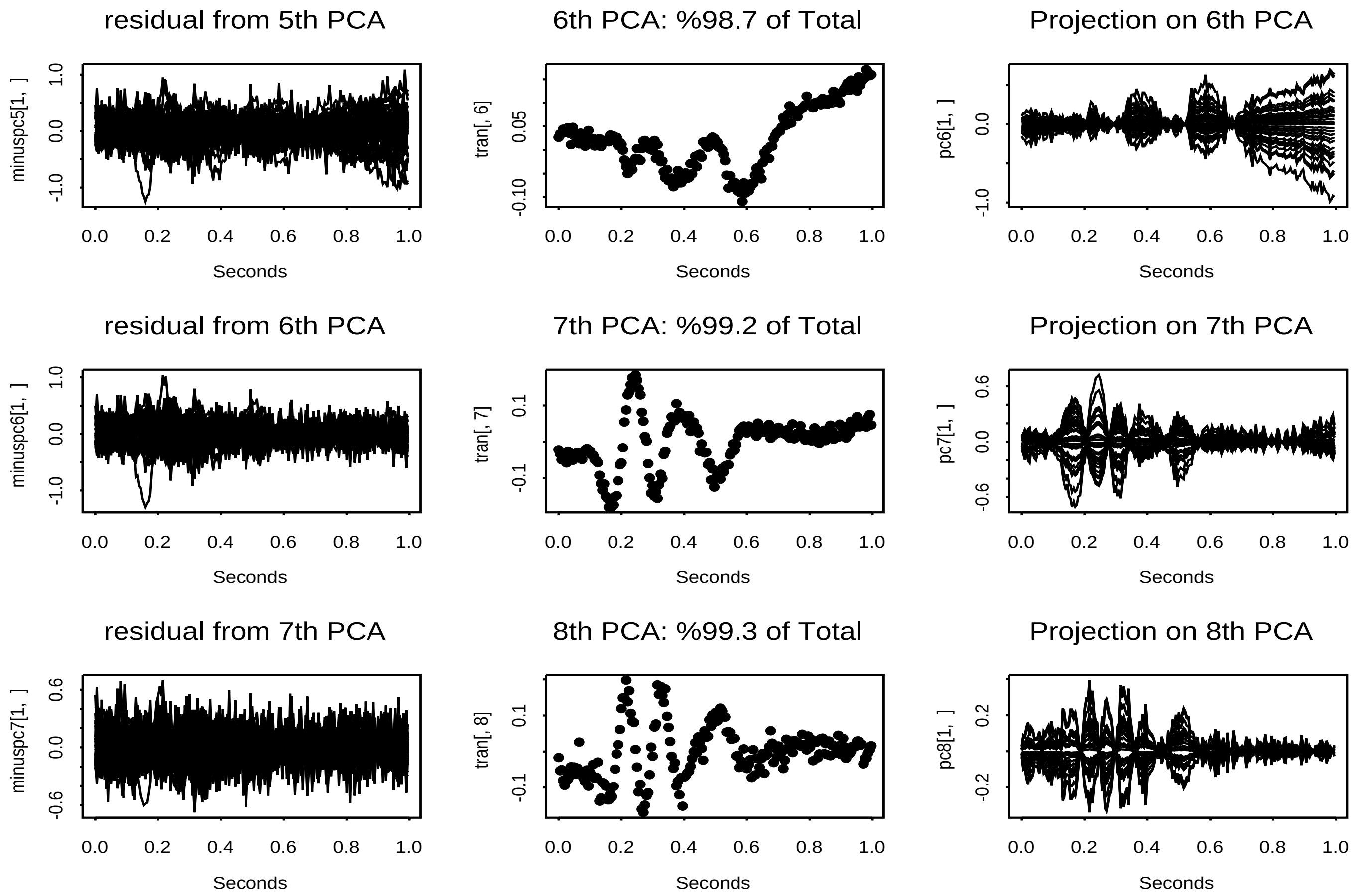


Figure 4: Third picture of PCA

Result of Method 1:

- principal component analysis gives us up to 7 or 8 curves, which are affected by the different location of the peak for each force record,
- does not give us a simple representation of this process.

Method 2:

(1) simple curve registration: subtracting the background force, adjusting curve by simple linear transformation:

$$s = \frac{t - t_0}{t_M - t_0} ,$$

where t_0 is the time of origin of the force impulse, t_M is the time of the maximum force,

(2) apply the principal component analysis on the adjusted curves.

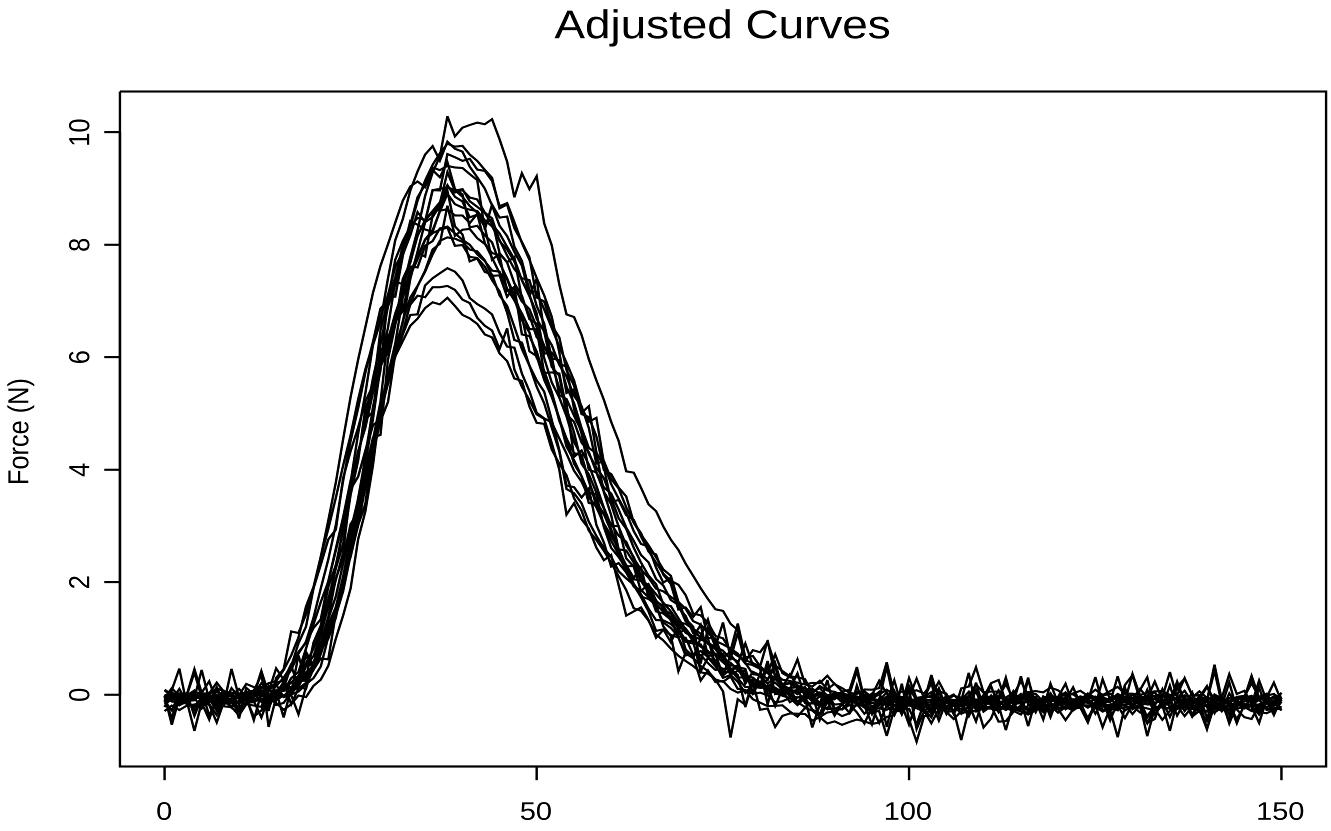


Figure 5:

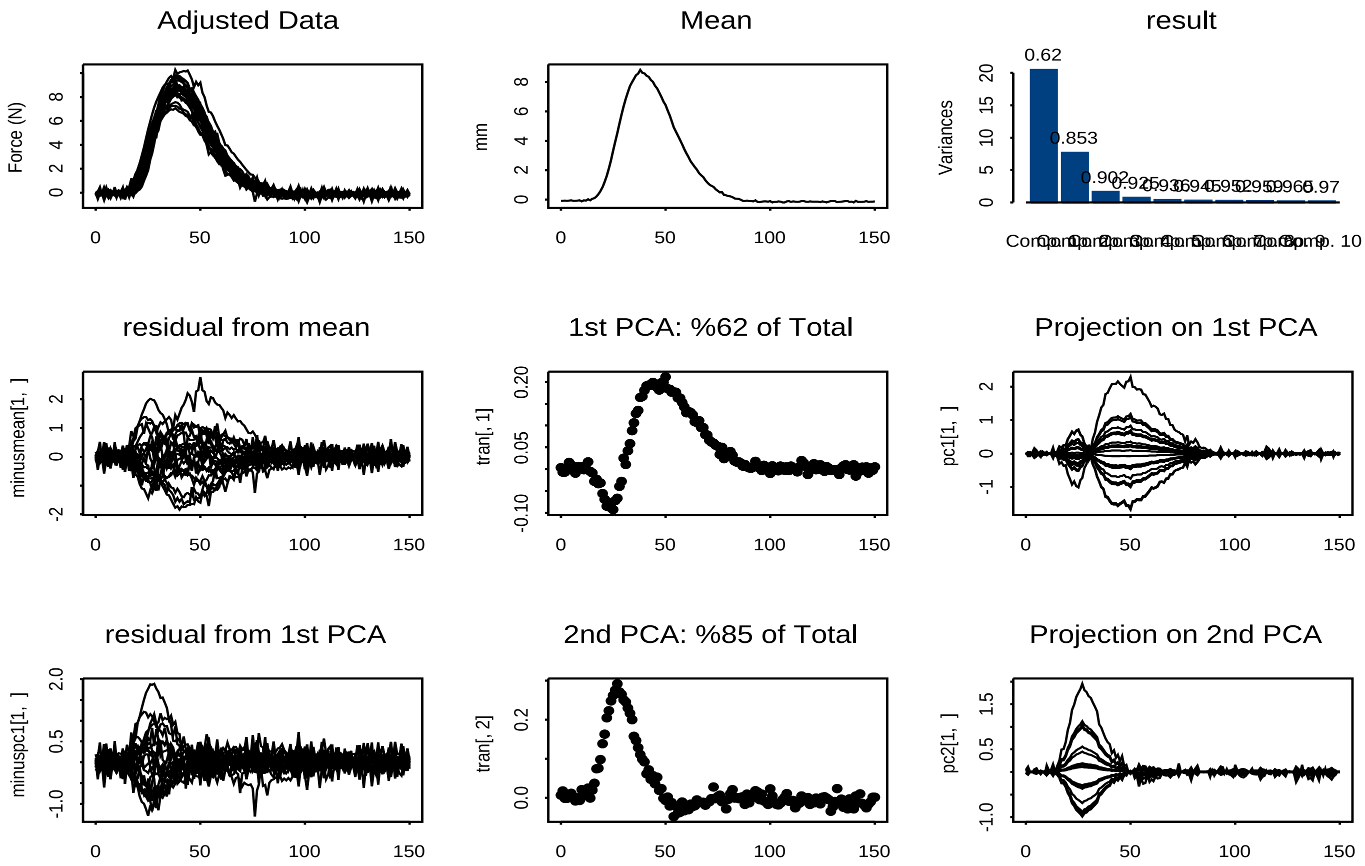
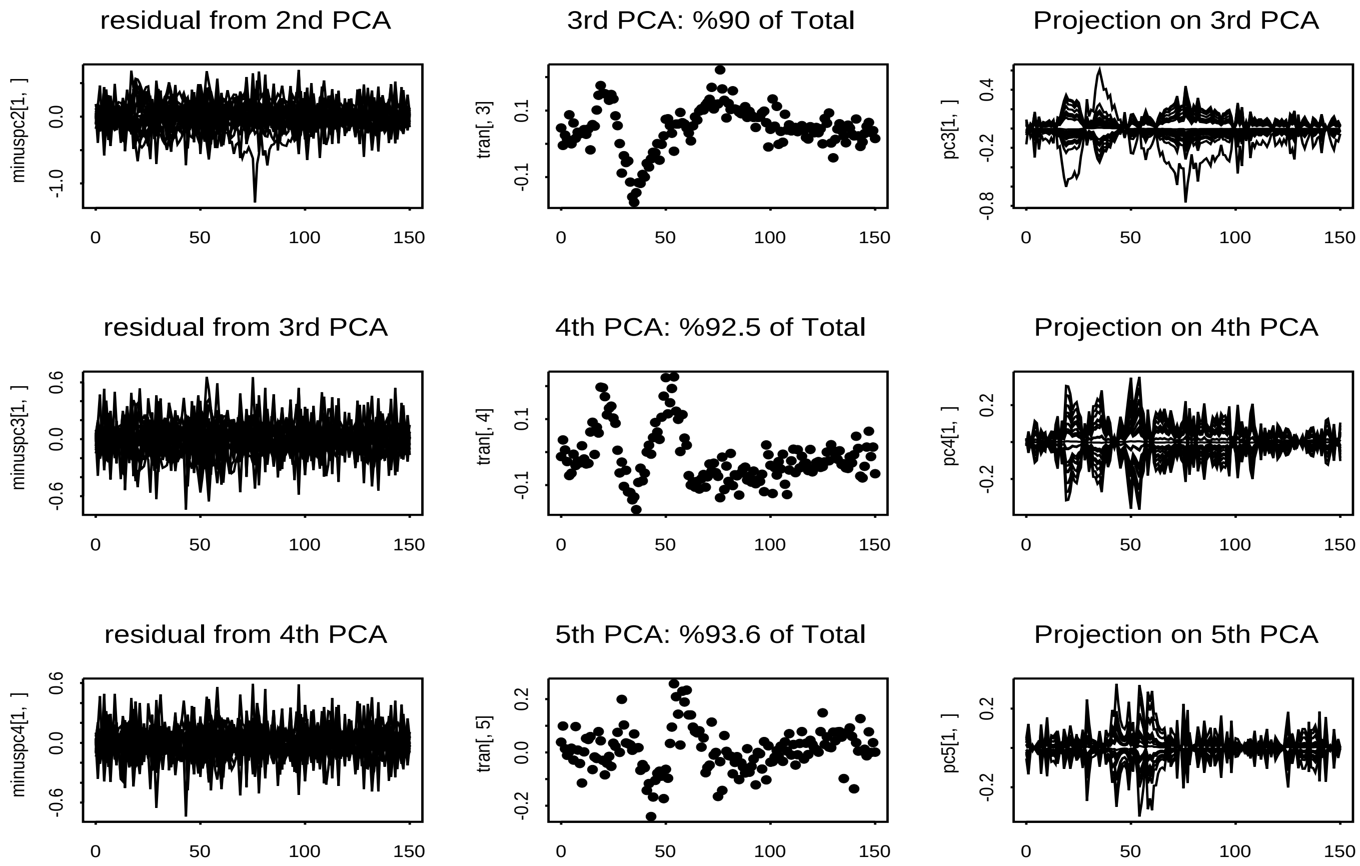
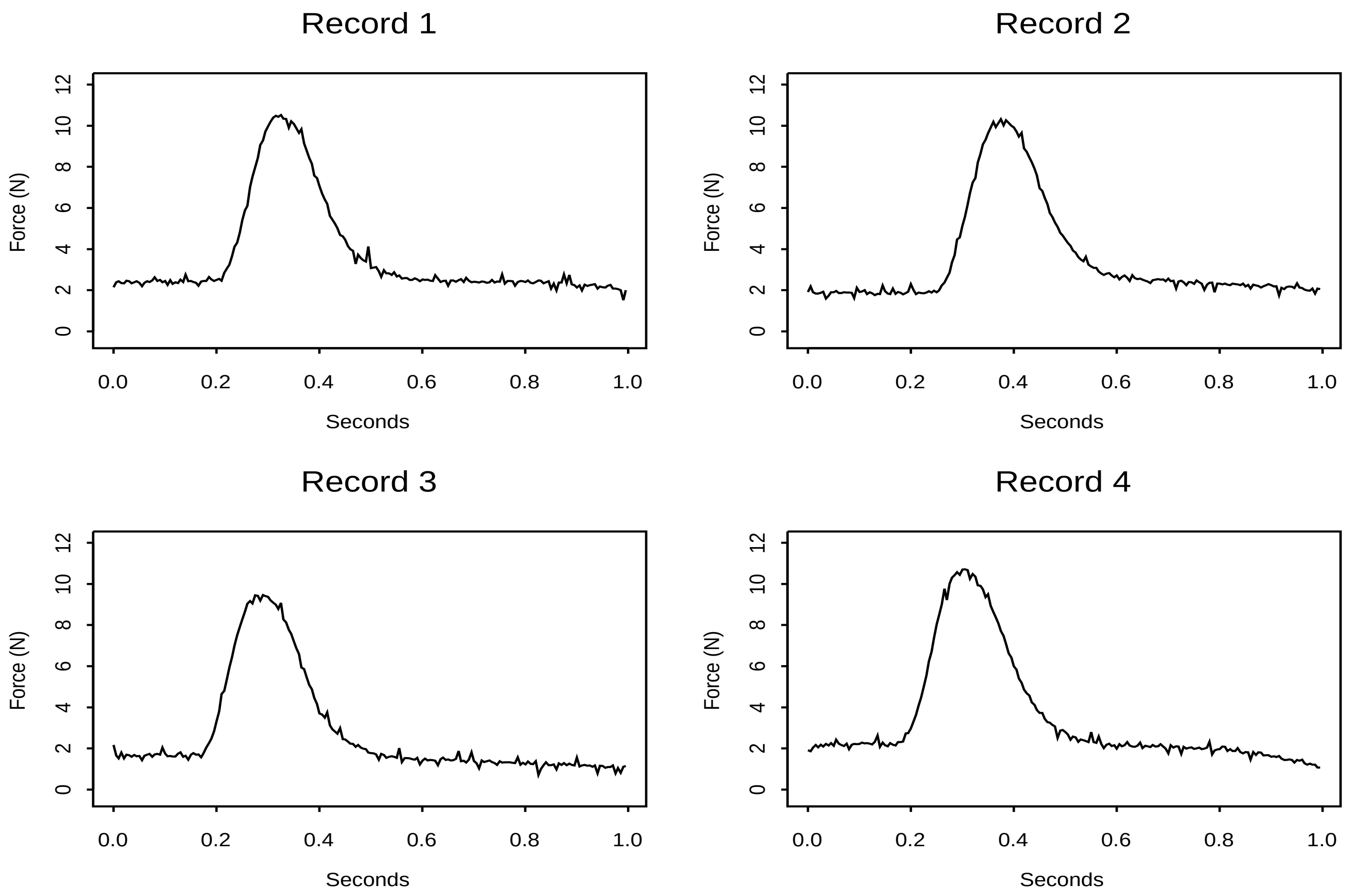

Adjusted Data
Force (N)
Mean
mm
result
Variances
0.62
0.853
residual from mean
minusmean[1,]
1st PCA: %62 of Total
tran[, 1]
Projection on 1st PCA
pc1[1,]
residual from 1st PCA
minuspc1[1,]
2nd PCA: %85 of Total
tran[, 2]
Projection on 2nd PCA
pc2[1,]

Figure 6:



residual from 2nd PCA
minuspc2[1,]
3rd PCA: %90 of Total
tran[, 3]
Projection on 3rd PCA
pc3[1,]
residual from 3rd PCA
minuspc3[1,]
4th PCA: %92.5 of Total
tran[, 4]
Projection on 4th PCA
pc4[1,]
residual from 4th PCA
minuspc4[1,]
5th PCA: %93.6 of Total
tran[, 5]
Projection on 5th PCA
pc5[1,]

Figure 7:



Record 1
Record 2
Record 3
Record 4
Force (N)
Seconds
0
2
4
6
8
10
12
0.0
0.2
0.4
0.6
0.8
1.0

Figure 8:

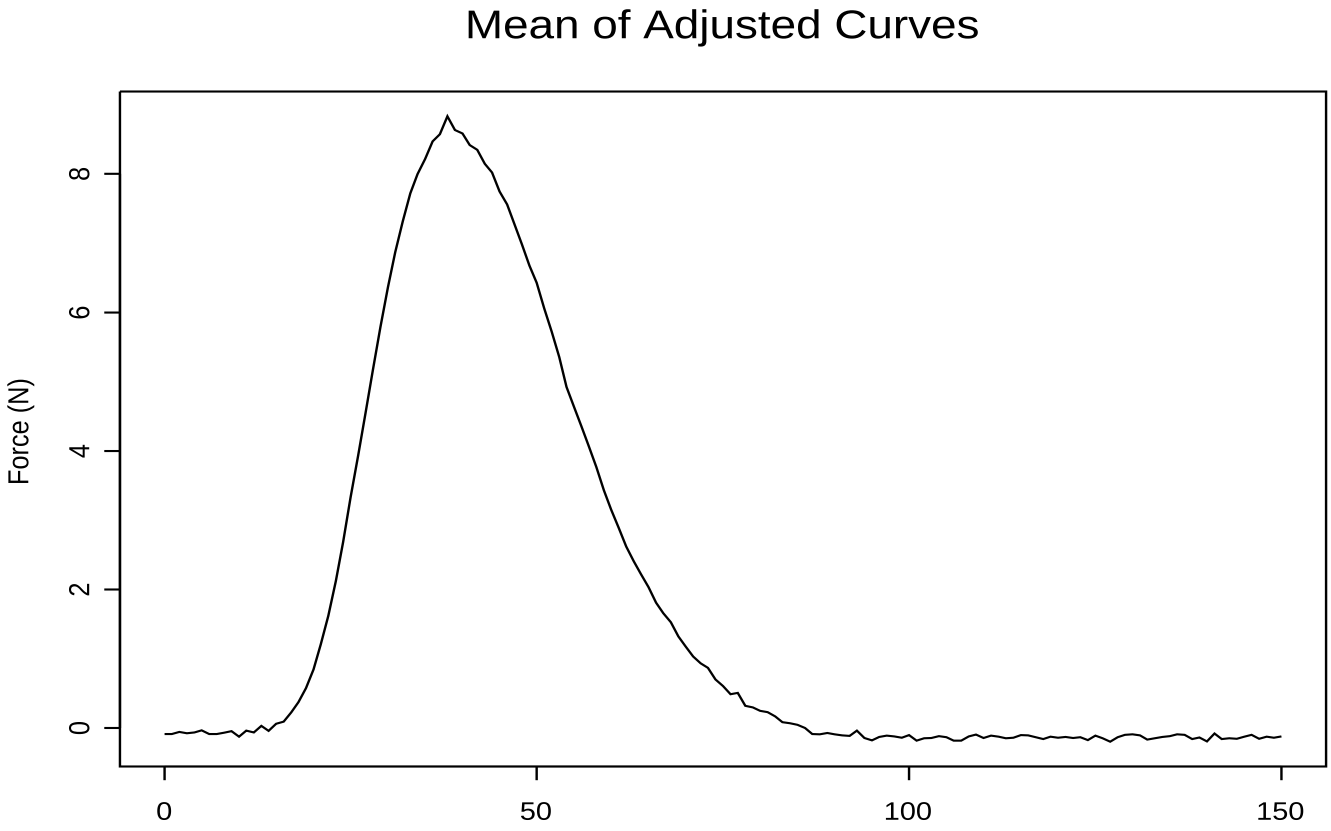


Figure 9:

Density of Log-normal:

$$p(t) = (2\pi)^{-1/2} t^{-1} \exp[-(\ln t - \ln \mu)^2/(2\sigma^2)] \ ,$$

where μ and σ are location and scale parameters, respectively. Or

$$p(t) = (2\pi)^{-1/2} \exp[-\ln \mu + \sigma^2/2] \exp[-(\ln t - (\ln \mu - \sigma^2))^2/(2\sigma^2)] \ .$$

the maximum of this function, or the modal value, is $t = \ln \mu - \sigma^2$.

This formulation suggests the model:

$$f(t) - f_0 = C \exp\{-[\ln(t - t_0) - \ln(t_M - t_0)]^2/(2\sigma^2)\} \ , t > t_0 \qquad (1)$$

where $f(t)$ is the force exerted at time t, f_0 is the targeted background force, C is the maximum force, t_0 is the time of origin of the force impulse, t_M is the time of the maximum force, and σ is a shape parameter.

Express the above model by

$$\ln f(t) = -b_0 - b_1 \ln(t - t_0) - b_2 \ln^2(t - t_0) \ ,$$

The quadratic approximation to log force can also be used to define a common origin and scale. The linear transformation of time

$$s = \frac{t - t_0}{t_M - t_0} \ ,$$

where t_M is the time of maximum force, implies that log force as a function of $\ln s$ is symmetric about $s = 1$, and hence that the model simplifies to

$$\ln f(s) = -b_0^* - b_2^* \ln^2 s$$

and

$$f(s) = \exp[-b_0^* - b_2^* \ln^2 s] = C \exp[-\ln^2 s/(2\sigma^2)] \ , \qquad (2)$$

where $C = e^{-b_0^*}$ and $\sigma^2 = (2b_2^*)^{-1}$.

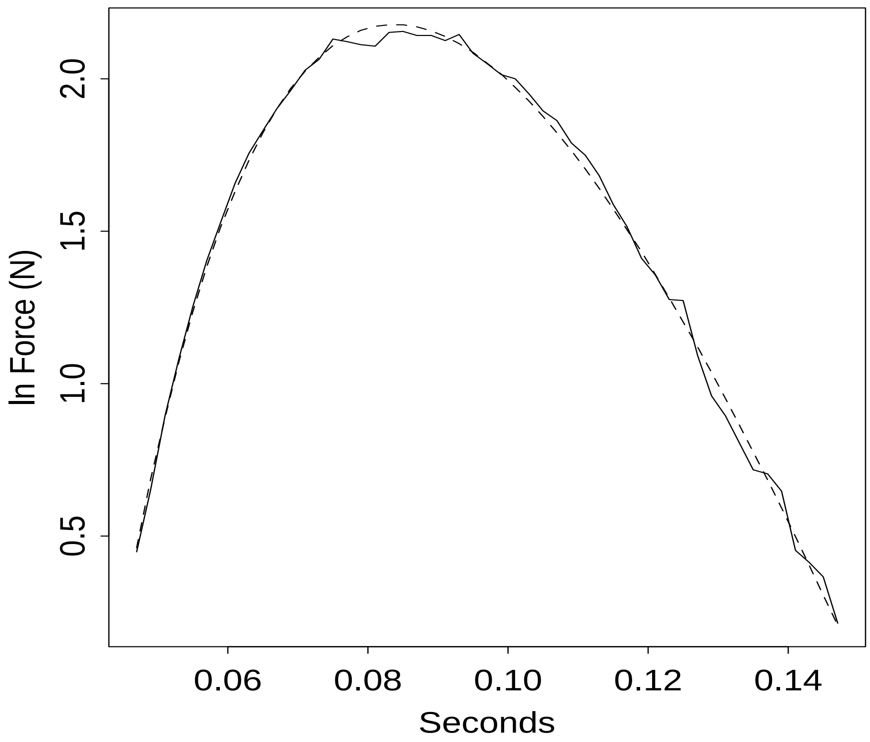


Figure 10:

A useful descriptor of the total response effort:

$$
\begin{aligned}
E &= (t_M - t_0) \int_0^\infty f(s) ds \\
&= (t_M - t_0) \sqrt{2\pi} C \sigma e^{\sigma^2/2} \,, \qquad (3)
\end{aligned}
$$

It is related:

- time to maximum force $t_M - t_0$,
- maximal force C
- the shape-dependent function $\sigma \exp(\sigma^2/2)$

Table 1: Summary Statistics for all Pinch Force Function Descriptors (after subtracting baseline force)

	$t_M - t_0$	C	σ	E
Mean	0.084	8.96	0.286	0.549
Std. Deviation	0.017	0.857	0.045	0.059
Coeff. Variation	0.197	0.096	0.157	0.107

Table 2: Correlations Among Pinch Force Function Descriptors

	$t_M - t_0$	C	σ	E
$t_M - t_0$	1.00	-0.31	-0.97	-0.24
C	-0.31	1.00	0.34	0.95
σ	-0.97	0.34	1.00	0.32
E	-0.24	0.95	0.32	1.00

.

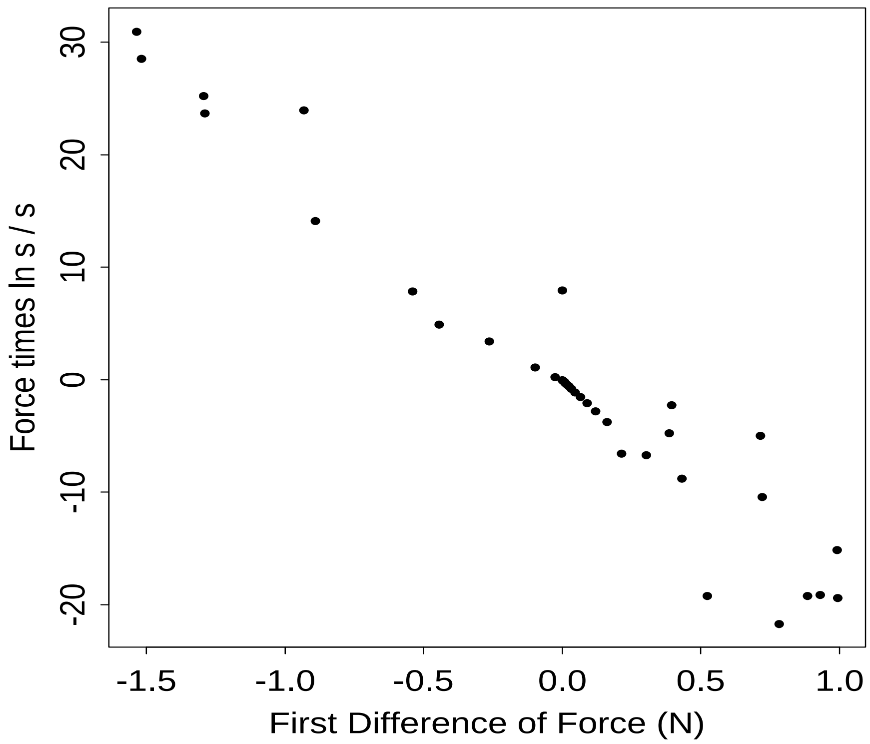
Model (2)

$$\ln f(s) = -b_0^* - b_2^* \ln^2 s$$

can also be expressed as a differential equation:

$$\frac{df}{ds} = -\frac{1}{\sigma^2}\frac{\ln s}{s} f(s) \tag{4}$$

Notice: $-1/\sigma^2$ is the slope in the linear relationship between f' and $(f \ln s)/s$, and therefore that σ^2 is the slope of the normal to this line.



Force times ln s / s
30
20
10
0
-10
-20
-1.5
-1.0
-0.5
0.0
0.5
1.0
First Difference of Force (N)

Figure 11:

The goal is to fit nonparametrically a record $(s_i, y_i), i = 1, \ldots, n$, where y_i is the observed force at sampling point i at transformed time s_i, by a more general smoothing function

$$h(s) = f(s) + e(s) \ .$$

If define a linear differential operator

$$(Lf)(s) = \frac{1}{\sigma^2}\frac{\ln s}{s} f(s) + \frac{df}{ds}$$

or

$$L = \frac{1}{\sigma^2}\frac{\ln s}{s} I + D$$

therefore $Lf = 0$ for any f in model

$$\ln f(s) = -b_0^* - b_2^* \ln^2 s,$$

and to uniquely identify e by constraint

$$\int_0^\infty e(s) ds = 0$$

Use the smoothing spline criterrion

$$\begin{aligned} Q &= \sum_{i=1}^{n} [y_i - h(s_i)]^2 + \lambda \int (Lh)^2(s) ds \\ &= \sum_{i=1}^{n} [y_i - h(s_i)]^2 + \lambda \int (Le)^2(s) ds \end{aligned}$$

The first term is a least squares error term, and the second term measures departure from smoothness in term of the size of the function $Lh = L(f + e) = Le$.

By (Wahba, 1990), the solution h to the spline smoothing problem for a fixed value of λ, minimizing

$$Q = \sum_{i=1}^{n} [y_i - h(s_i)]^2 + \lambda \int (Lh)^2(s) ds$$

can be expressed in the form:

$$h(s) = \sum_{i=1}^{m} d_i f_i(s) + \sum_{j=1}^{n} c_j k(s_j, s)$$

where the functions f_i span the space of functions satisfying $Lf = 0$, and where the bivariate function k is the reproducing kernel for the Hilbert space of functions e satisfying $\int e = 0$ and with the inner product

$$(e_j, e_k) = \int (Le_j)(s)(Le_k)(s) ds \ .$$

The data-driven generalized cross-validation technique by Wahba (1990) was applied to choose an appropriate value for the smoothing parameter λ

Let vector $\mathbf{f}$ contain the values $f(s_j)$, $\mathbf{y}$ contain y_j, and matrix $\mathbf{K}$ contain the values $k(s_i, s_j), i, j = 1, \ldots, n$. Then the coefficient d and coefficient vector $\mathbf{c}$ defining the smoothing spline h satisfy the linear equations

$$\mathbf{Mc} + \mathbf{f}d = \mathbf{y}$$

$$\mathbf{f}'\mathbf{c} = 0$$

from which we have

$$d = \mathbf{f}'\mathbf{M}^{-1}\mathbf{y}/\mathbf{f}'\mathbf{M}^{-1}\mathbf{f}$$

and

$$\mathbf{c} = \mathbf{M}^{-1}(\mathbf{y} - \mathbf{f}d)$$

where $\mathbf{M} = \mathbf{K} + n\lambda\mathbf{I}$.

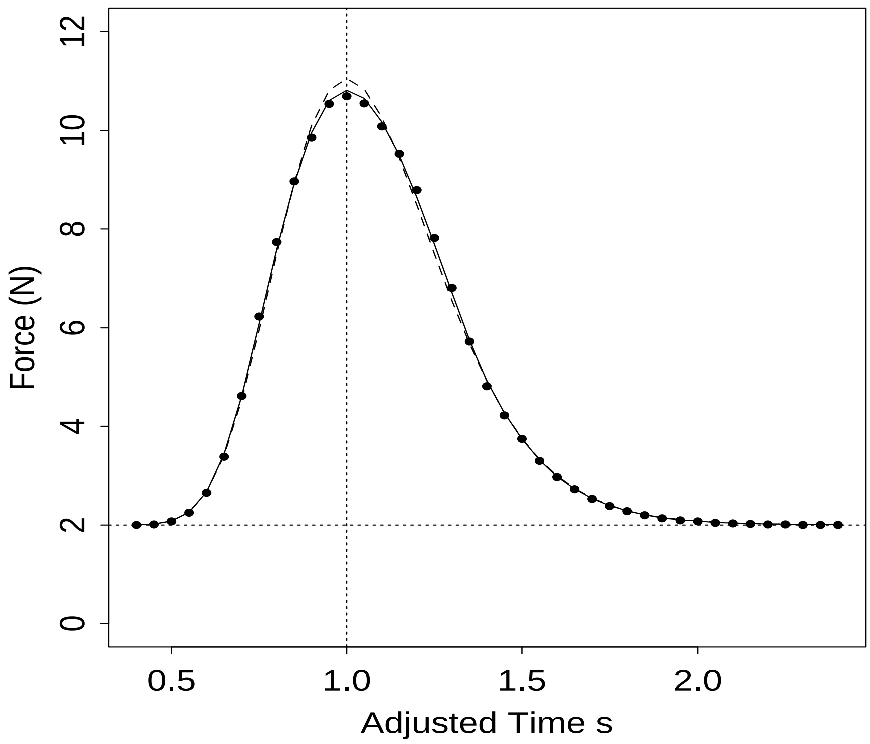


Figure 12: Smoothing spline fit h by solid line, dot points are observed forces, dash line is the log-normal model component f

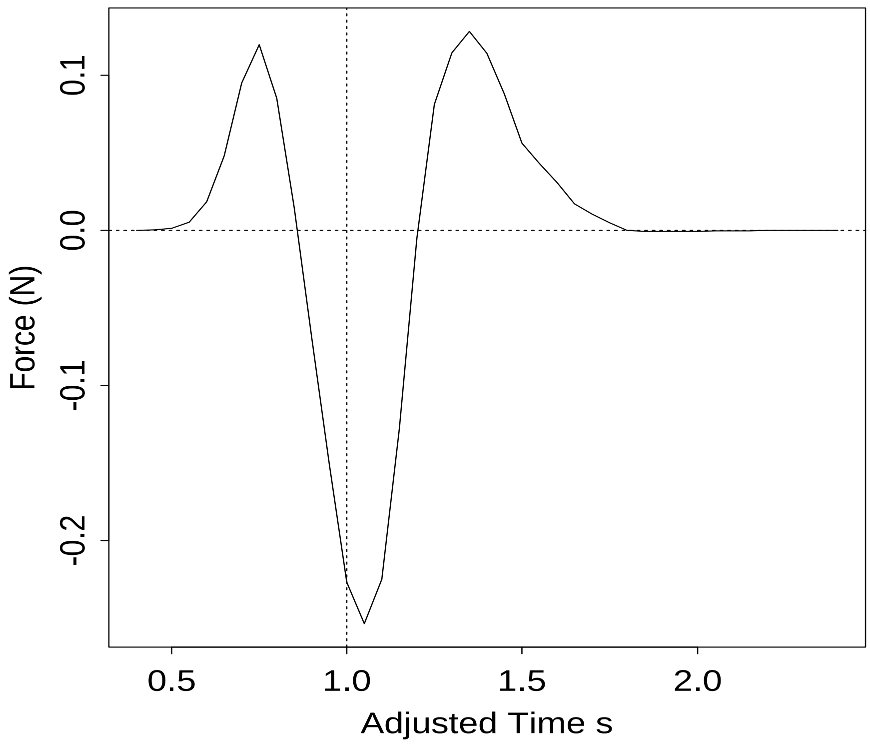


Figure 13: The average residual function $\bar{e}$ for a set of force impulses where target was 10N and the base-line force was 2N.

For multivariate statistics, let x and y be vectors of p-dimension Euclidean space, and the inner product is the Euclidean inner product

$$\langle x, y \rangle = \sum_{i=1}^{p} x_i y_i$$

Now consider x as function with values $x(t)$, where argument t takes values on some real interval l, e.g. $l = (-\infty, +\infty)$, $l = [0, +\infty)$, or $l = [0, T]$. Therefore the index i taking p possible values has been replaced by the index t taking an infinity of values. The corresponding inner product is

$$\langle x, y \rangle = \int_l x(t) y(t) dt.$$

For multivariate analysis, the covariance matrix is

$$\Sigma = \frac{1}{n-1}\mathbf{X}'\mathbf{X}$$

Finding the first principal component vector means finding ξ such that

$$\mathrm{m}ax_{\xi'\xi=1}\xi'\Sigma\xi$$

This maximization problem is solved by finding the solution with largest eigenvalue λ_1 of the eigenvector problem or eigenequation

$$\Sigma\xi = \lambda_1\xi$$

When $x(t)$ is a function of t instead of a p-dimension vector, the functional PCA leads to the eigenequation

$$\int v(s,t)\xi(t)dt = \langle v(s,\cdot), \xi \rangle = \lambda_1 \xi(s) \tag{5}$$

where the covariance function v is

$$v(s,t) = \frac{1}{n-1} \sum_{i=1}^{n} x_i(s) x_i(t)$$

The left side of (5) is an **integral transform** Σ of the function ξ with the kernel of the transform v defined by

$$\Sigma\xi = \int v(\cdot, t)\xi(t)dt.$$

Further, we still can express it as

$$\Sigma\xi = \lambda_1 \xi$$

The motivation for the choice of fitting criterion

$$Q = \sum_{i=1}^{n} [y_i - h(s_i)]^2 + \lambda \int (Lh)^2(s)ds$$

was that the transformed functions Le might be of interest. Applying the functional principal components analysis at Le level rather than to e and v directly, consequently, it is the eigenequation

$$\int (Lv(s,\cdot))(w)(L\xi)(w)dw = \lambda_1 \xi(s)$$

that is solved for eigenvalue λ_1 and eigenfunction ξ. Any two eigenfunctions are required to satisfy the orthonormality constraint

$$\int (L\xi_j)(s)(L\xi_k)(s)ds = \delta_{jk} \ ,$$

where $\delta_{jk} = 1$ if $j = k$ and 0 otherwise.

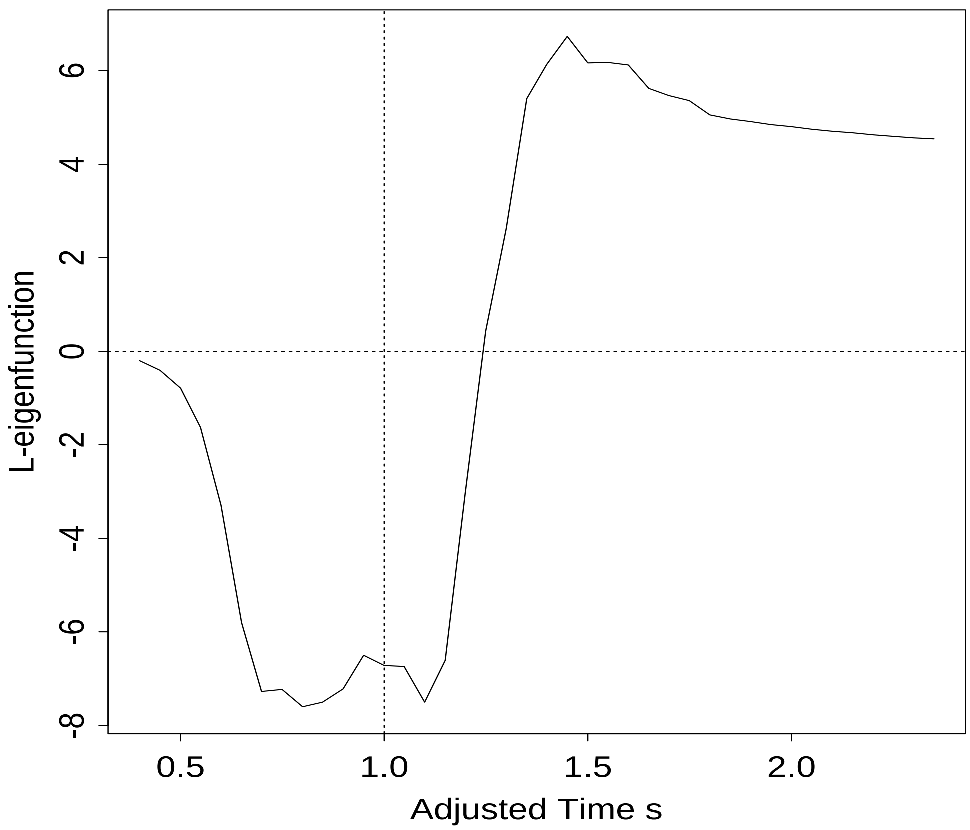


Figure 14: The result of applying differential operator L to the first eigenfunction of the variance-covariance function for the force impulse data.

Discussion and Conclusion The pinch force system can be described to an impressive level of accuracy by the differential equation

$$\frac{df}{ds} + \frac{1}{\sigma^2}\frac{\ln s}{s} f(s) = u(s) \ ,$$

where

$$u(s) = \begin{cases} U_1 < 0 \ , & s \leq 1.2 \\ U_2 > 0 \ , & s \geq 1.2 \end{cases}$$

and $s = (t - t_0)/(t_M - t_0)$.