Analysis of phycobilisomes mobility on thylakoid membrane: Part II. A single parameter estimation problem

Radek Kaňa, Ctirad Matonoha, Štěpán Papáček

Institute of Microbiology, AS CR, Opatovický mlýn, Třeboň
Institute of Computer Science, AS CR, Prague
Institute of Physical Biology, University of South Bohemia, Nové Hrady

9. seminář Matematika na vysokých školách

Herbertov - Horní Mlýn, 12. - 14. září 2011

This work was supported by the Grant Agency of the Czech Republic through the research grant No. 102/05/0011,
the institutional research plan No. AV0Z10300504,
and the project „Jihočeské výzkumné centrum akvakultury a biodiverzity hydrocenóz“ (CENAKVA), OP VaVpI.

The object of our attention is the following one-dimensional form of diffusion-reaction equation with spatial coordinate $x \in [0, 1]$, the diffusion coefficient $p > 0$, the time $\tau \geq \tau_0$, and the normalized phycobilisome particle concentration y:

$$\frac{\partial y}{\partial \tau} - p \frac{\partial^2 y}{\partial x^2} = -\alpha k_S y. \tag{1}$$

Here k_S is a reaction term and $\alpha = \alpha(D_0) > 0$ is a constant. The initial condition and time varying Dirichlet boundary conditions are:

$$y(x, \tau_0) = f(x), \quad x \in [0, 1], \tag{2}$$

$$y(0, \tau) = g_0(\tau), \quad y(1, \tau) = g_1(\tau), \quad \tau \geq \tau_0, \tag{3}$$

alternatively

$$\frac{\partial y}{\partial x}(0, \tau) = h_0(\tau), \quad \frac{\partial y}{\partial x}(1, \tau) = h_1(\tau), \quad \tau \geq \tau_0. \tag{4}$$

Based on FRAP experiments, we have a 2D matrix of dimension $(N+1, M+m+1)$ with prebleach and postbleach experimental values

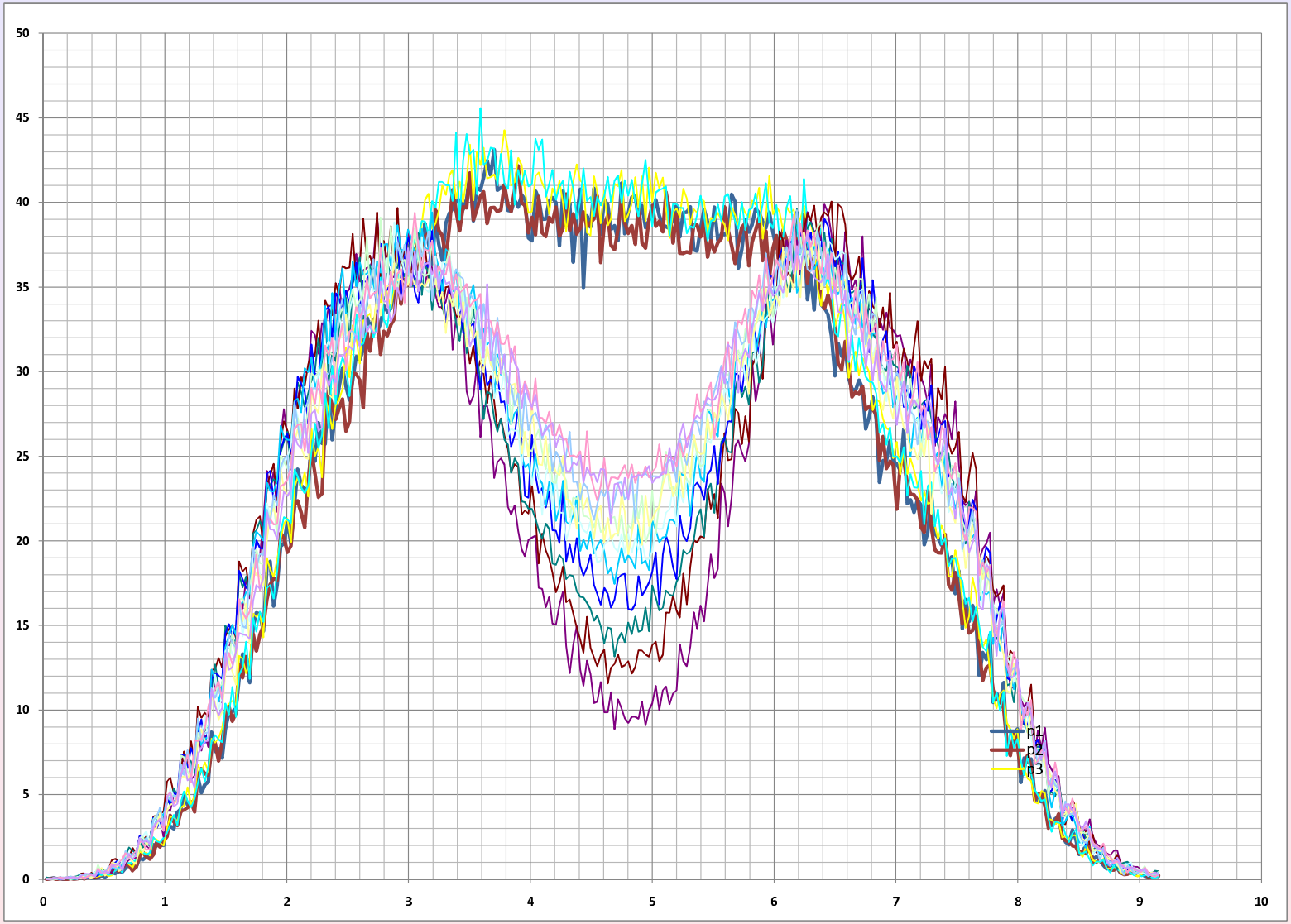
$$y_{exp}(x_i, \tau_j),\ i = 0 \ldots N,\ j = -m \ldots M,$$

with $x_0 = 0$ and $x_N = 1$, in fact characterizing the diffusion process, where

- $y_{exp}(x_i, \tau_0),\ i = 0 \ldots N$, is a vector for the IC,
- $y_{exp}(0, \tau_j),\ j = 0 \ldots M$, is a vector of the left Dirichlet BC,
- $y_{exp}(1, \tau_j),\ j = 0 \ldots M$, is a vector of the right Dirichlet BC,
- the BC are not yet implemented.

Recall that we have not a smooth function for the initial and boundary conditions. The forthcoming task is the analysis of measurement noise and its correct filtering.

Experimental values



Experimental values:

- space interval between first and last measurements: $[a, b]$
- length of space interval: $L = b - a$
- re-scaled dimensionless space interval: $x \in [0, 1]$
- re-scaled distance between two space measurements: $h = \frac{1}{N}$
- time interval between two measurements: T
- re-scaled dimensionless time interval: $\tau_t = \frac{TD_0}{L^2}$

We construct an objective function J representing the disparity between the experimental and simulated time-varying concentration profiles, and then within a suitable method we look for such a value p minimizing J. The usual form of an objective function is the sum of squared differences between the experimentally measured and numerically simulated time-varying concentration profiles:

$$J(p) = \sum_{j=0}^{M} \sum_{i=0}^{N} \left[y_{exp}(x_i, \tau_j) - y_{sim}(x_i, \tau_j) \right]^2 , \tag{5}$$

where $y_{sim}(x_i, \tau_j)$ are the simulated values resulting from the solution of problem (1)–(3). Firstly we started neglecting the reaction term (i.e. we put $k_S = 0$).

Minimizing J with respect to p represents a one-dimensional optimization problem. We have used a suitable optimization method from the so-called UFO system which generates a sequence of iterates $\{p_l,\ l \geq 0\}$ leading to a value p_* which minimizes J.

Usually, only postbleach data are considered. However, we can also take into account prebleach data. In this case, we compute the average value of them and now

- either average prebleach is subtracted from postbleach data,
- or postbleach data are divided by average prebleach.

To remove the noise in experimental values, we can further consider their smoothing by using the Fourier transformation.

In both the above cases we have „new“ experimental values y_{exp}.

In order to compute a function value $J(p_l)$ in (5) for a given p_l in the l-th iteration, we need to know both

- the *experimental* values $y_{exp}(x_i, \tau_j),\ i = 0 \dots N,\ j = 0 \dots M,$
- the *simulated* values $y_{sim}(x_i, \tau_j),\ i = 0 \dots N,\ j = 0 \dots M.$

It means that in each l-th iteration we need to solve the problem ($y_{sim} \equiv y$)

$$\frac{\partial y}{\partial \tau} - p_l \frac{\partial^2 y}{\partial x^2} = 0 \tag{6}$$

with the initial and boundary conditions defined by the experimental data:

$$y(x, \tau_0) = y_{exp}(x, \tau_0) \quad \text{for} \quad x \in [0, 1], \tag{7}$$

$$y(0, \tau) = y_{exp}(0, \tau), \quad y(1, \tau) = y_{exp}(1, \tau) \quad \text{for} \quad \tau \geq \tau_0. \tag{8}$$

Problem (6)-(8) for simulated data $y(x_i, \tau_j)$ was solved numerically using the finite difference scheme for uniformly distributed nodes with the space steplength Δh and the variable time steplength $\Delta\tau$:

1. The explicit scheme of order $\Delta\tau + \Delta h^2$:

$$y_{i,j} = \beta y_{i-1,j-1} + (1 - 2\beta) y_{i,j-1} + \beta y_{i+1,j-1}$$

2. The Crank-Nicholson implicit scheme of order $\Delta\tau^2 + \Delta h^2$:

$$-\frac{\beta}{2} y_{i-1,j} + (1+\beta) y_{i,j} - \frac{\beta}{2} y_{i+1,j} = \frac{\beta}{2} y_{i-1,j-1} + (1-\beta) y_{i,j-1} + \frac{\beta}{2} y_{i+1,j-1}$$

Here $\beta = \frac{\Delta\tau}{\Delta h^2}\, p$ and $y_{i,j} \equiv y(x_i, \tau_j)$ are the computed values in nodes. Recall that for the explicit scheme the condition $\beta \leq 1/2$ must hold.

Steplengths used in the numerical scheme:

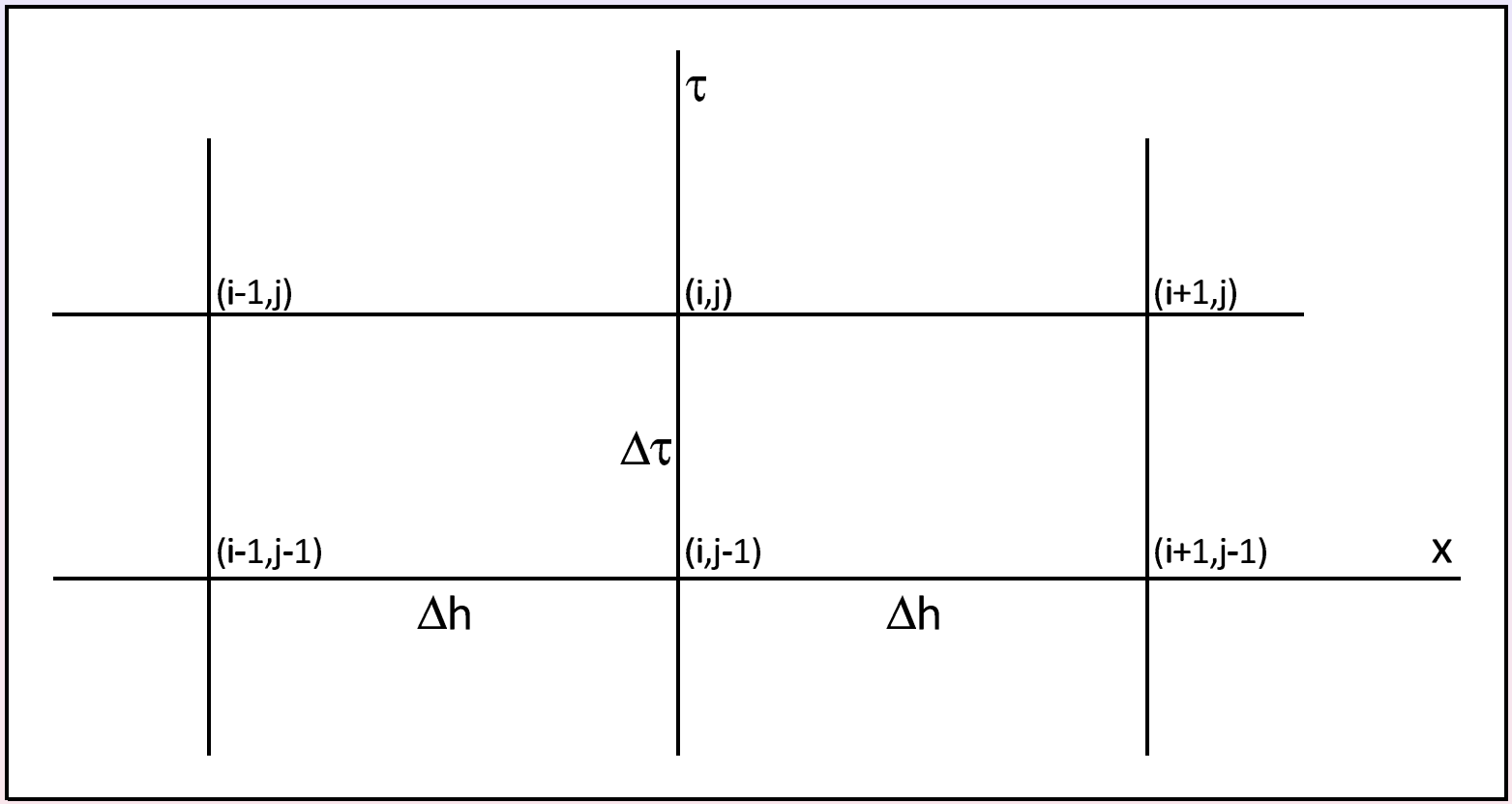
- Space steplength: $\Delta h = 1/N$ or $\Delta h = 1/(\kappa_s N),\ \kappa_s \in \mathbb{N}$
- Time steplength: $\Delta\tau$ should be ideally of the same order as Δh^2 (or Δh in the CN scheme) and in the explicit scheme has to fulfill the relation $\Delta\tau \leq \frac{\Delta h^2}{2p}$.

In order to get from the $(j-1)$-th time row to the j-th, we need to perform

$$\kappa_t = \frac{TD_0}{L^2\Delta\tau}$$

substeps, where $\kappa_t \in \mathbb{N}$ has to be an integer depending on $\Delta\tau$.

Differential scheme



Taking into account the biological reality residing in possible time dependence of phycobilins diffusivity, for the minimization problem

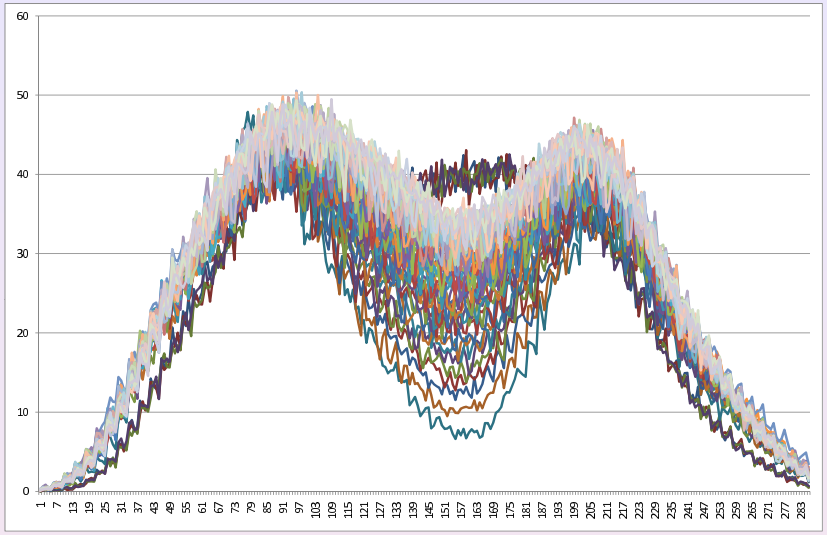
$$\min_p J(p) = \min_p \sum_{j=0}^{M} \sum_{i=0}^{N} \left[y_{exp}(x_i, \tau_j) - y_{sim}(x_i, \tau_j) \right]^2 \tag{9}$$

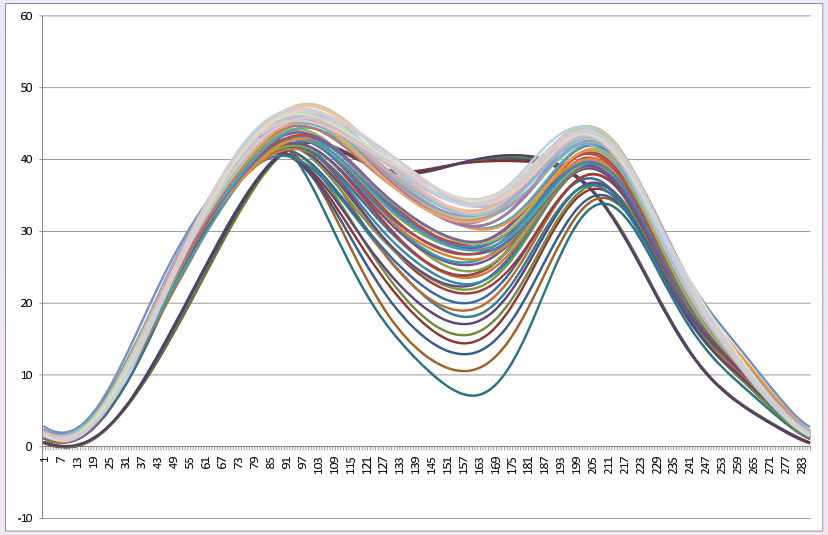
we further consider two cases:

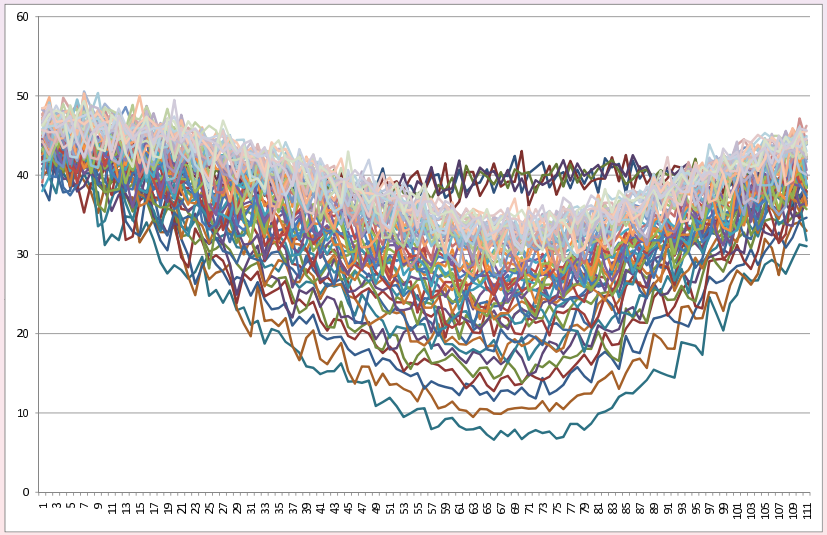
1. We can take both sums for i and j in (9) together. In this case, the scalar p is a result of minimization problem for J.
2. We can consider each j-th time row separately. In this case, the M solutions $p^{(1)}, ..., p^{(M)}$ correspond to each minimization problem for fixed j in the sum (9) and we have a 'dynamics' of diffusivity p evolution.

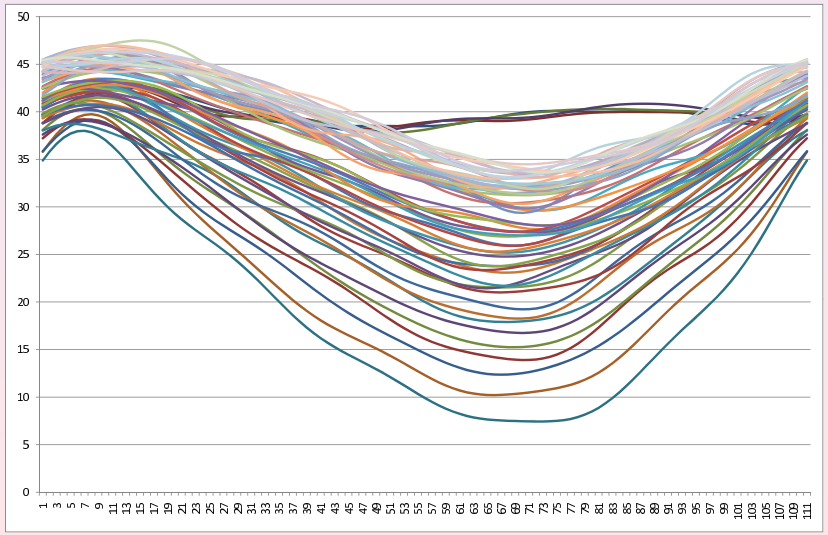
- N – počet vnitřních uzlů včetně krajních
- H – prostorový krok (v metrech)
- M – počet časových vrstev včetne počáteční
- T – časový krok (v sekundách)
- D0 – fyzikální difuzní koeficient D_0 (v m^2/s)
- SCHEMA – 1: explicitní, 2: implicitní
- ZHLAZENI – 0: žádné, 1..9: počet členů FFT
- PREBLEACH – 0: ne, > 0: kolik sloupců tvoří prebleach, udělá se průměr, kterým se odečtou následná data, < 0: dtto, vydělí
- NORMALIZACE – 0: ne, 1: ano: $J := J/(N-1)$
- PRIMA – 0: pouze invertní úloha, 1: pouze přímá úloha
- A B – pracuje se s daty od $A >= 1$ do $B <= N$
- HOT – 1: pro hot starty, 0: pro warm starty
- XINIT – počáteční p_0 (jen jednou pro HOT=1, vždy pro HOT=0)
- XMAX – velikost oblasti, kde se hledá iterace $p_l \to p_{l+1}$
- TOLG – tolerance na řešení $p_* = p^{(j)}$
- Y-SCHEMA – výpis vstupních dat y, se kterými se dělá schema

Data: A=1, B=286: FFT=0, 5; A=90, B=200: FFT=0, 5

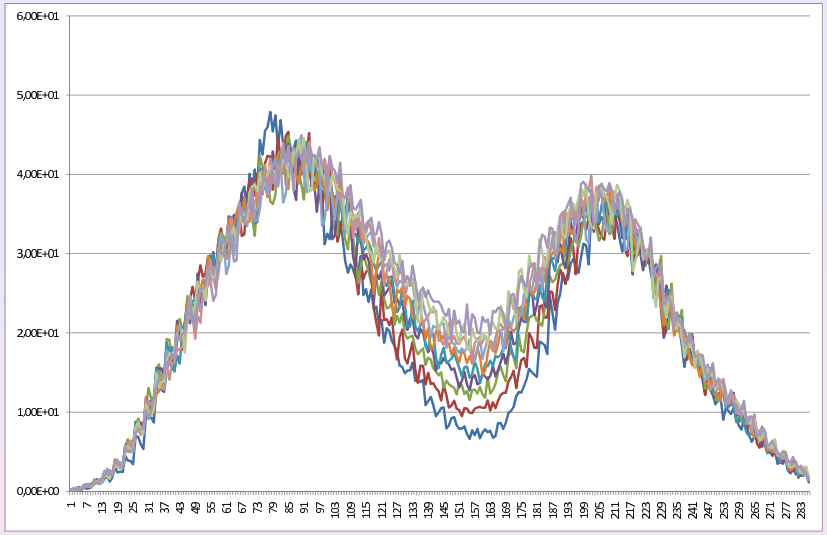


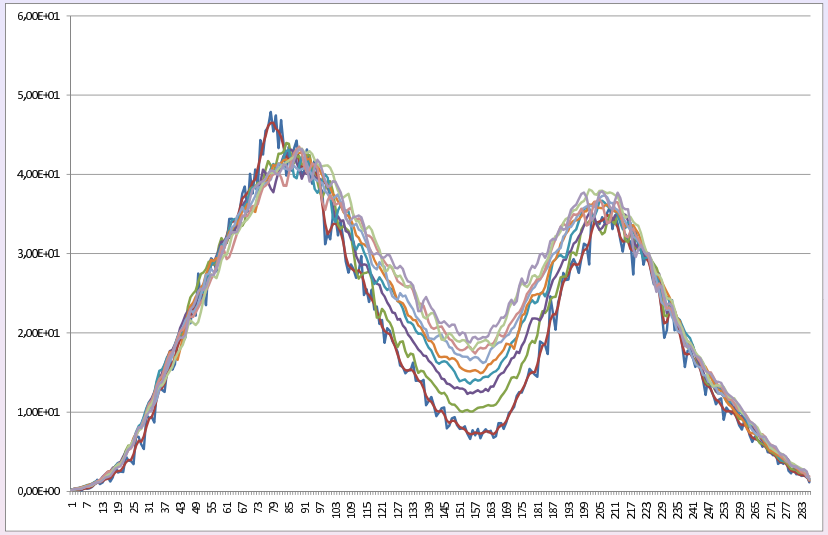


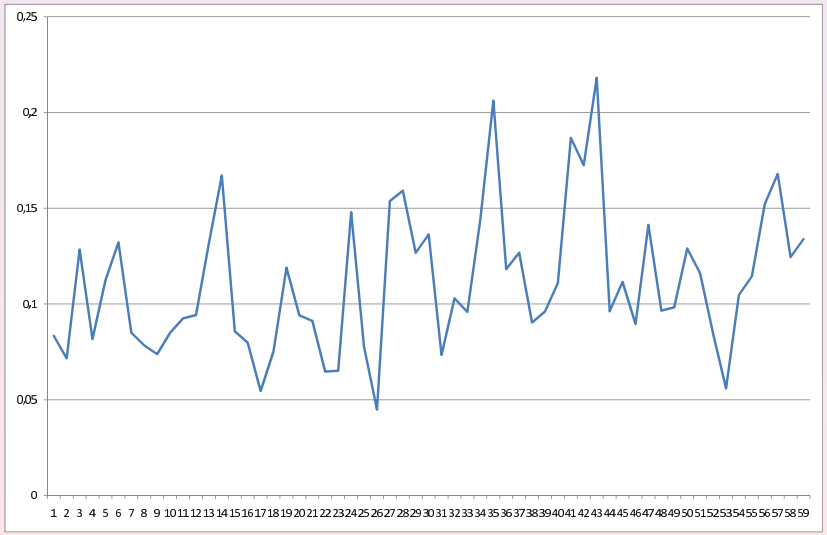


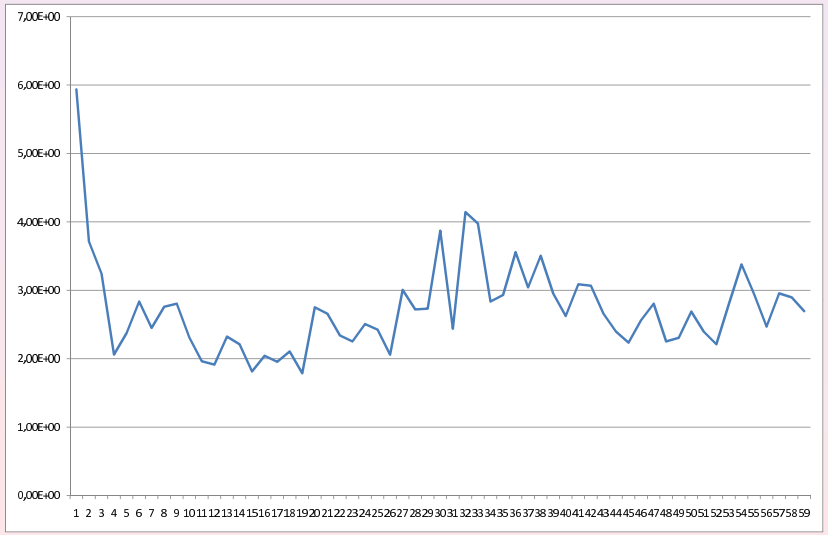


Data01: FFT=0, prebleach=0: y-schema, y-solution, p, J

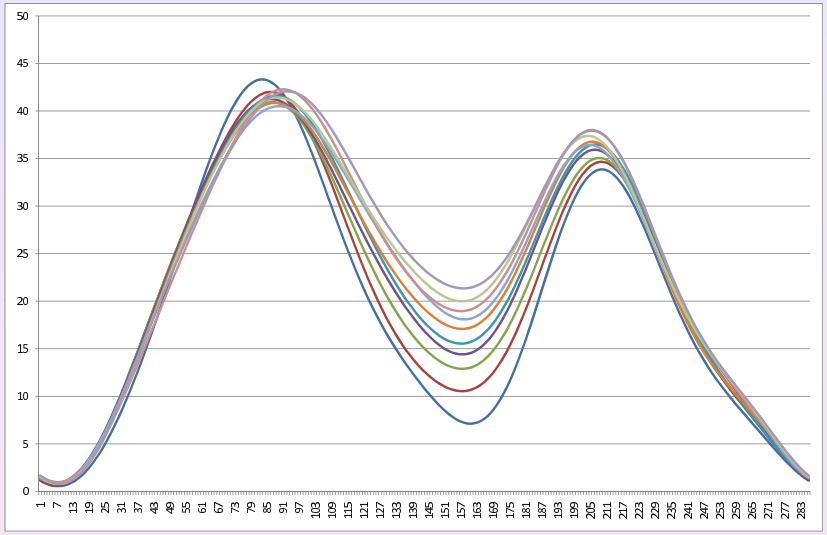


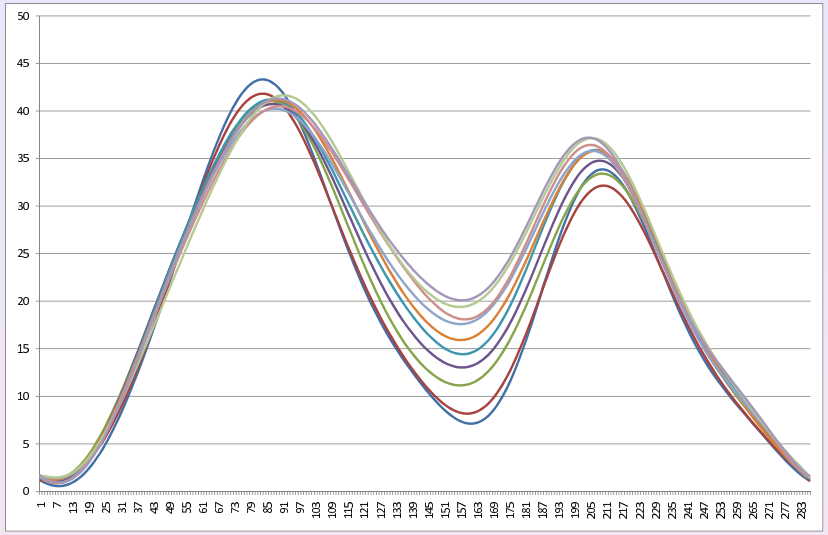


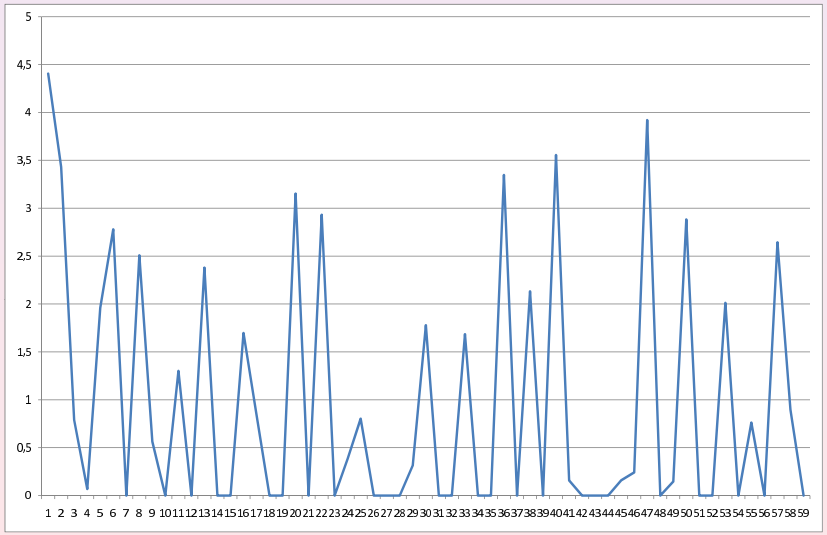


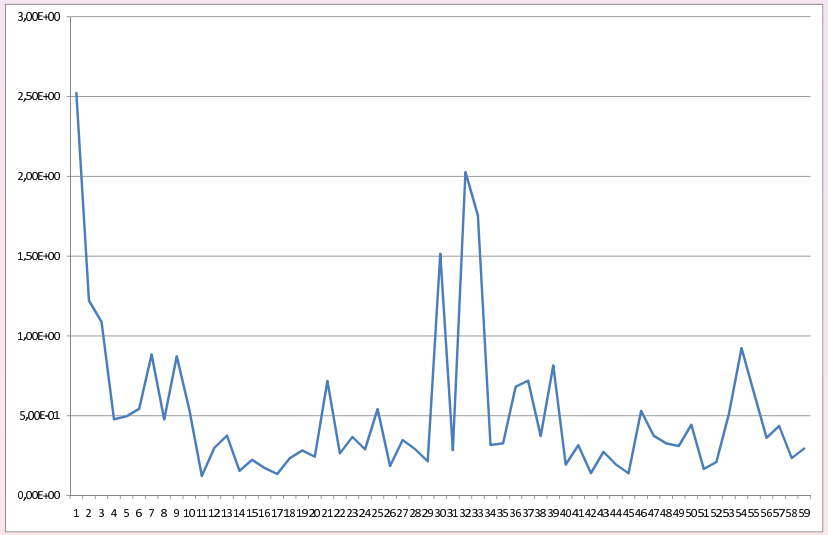


Data02: FFT=5, prebleach=0: y-schema, y-solution, p, J

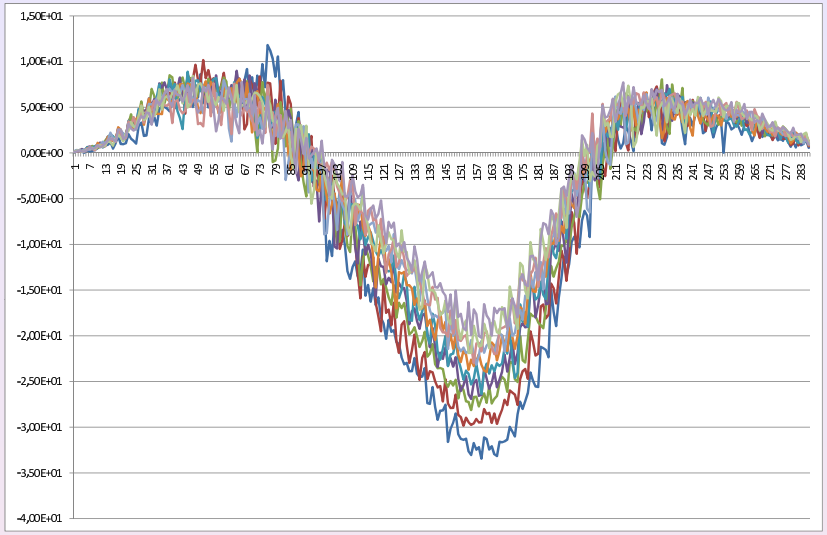


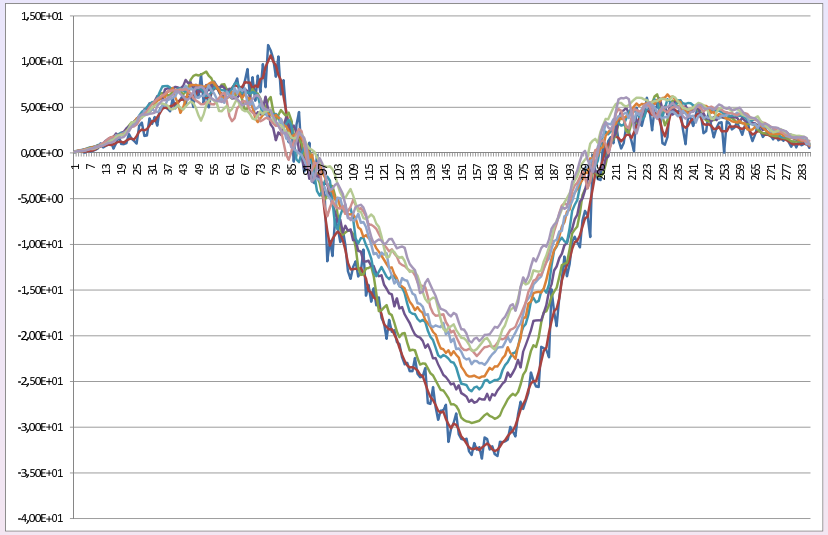


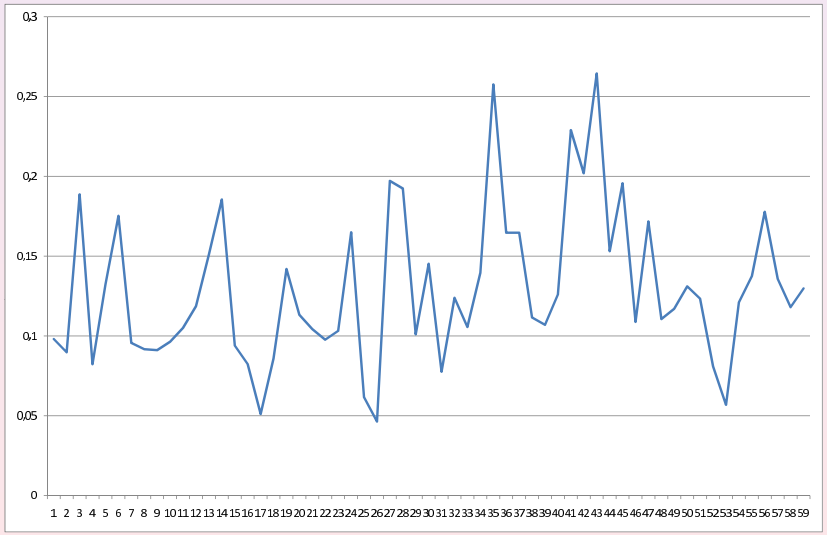


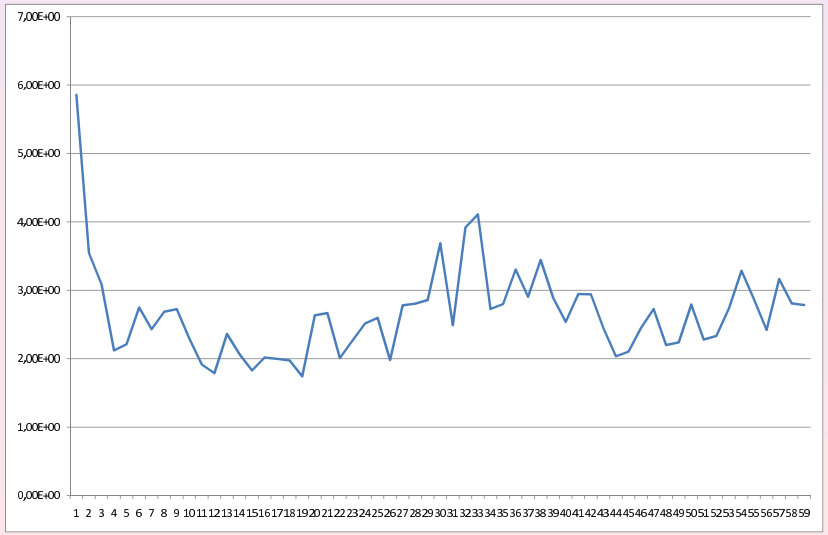


Data03: FFT=0, prebleach=4: y-schema, y-solution, p, J

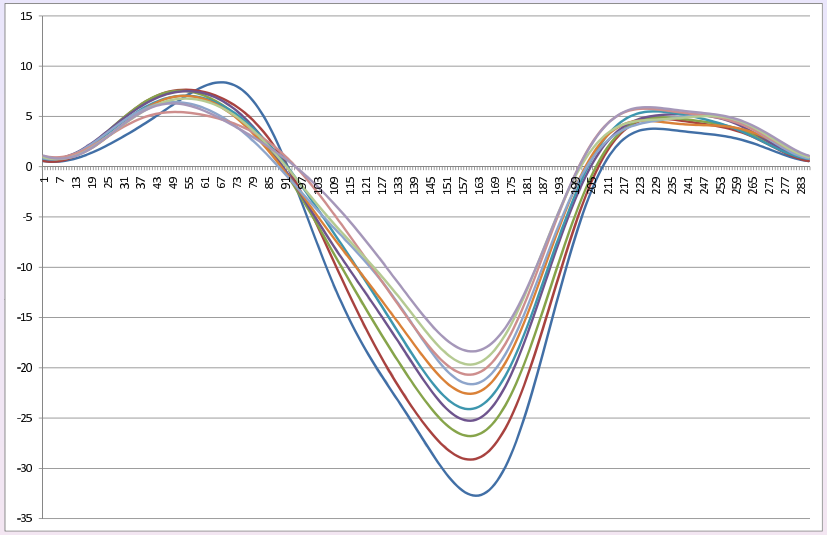


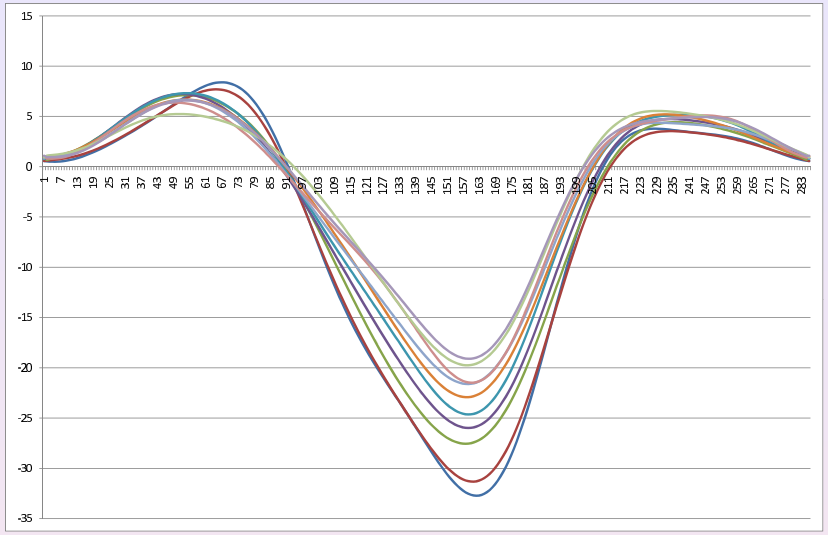


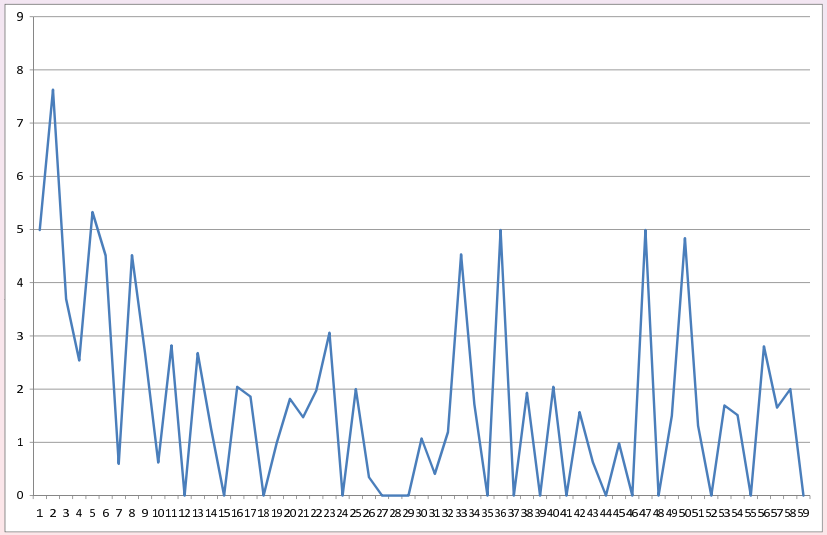


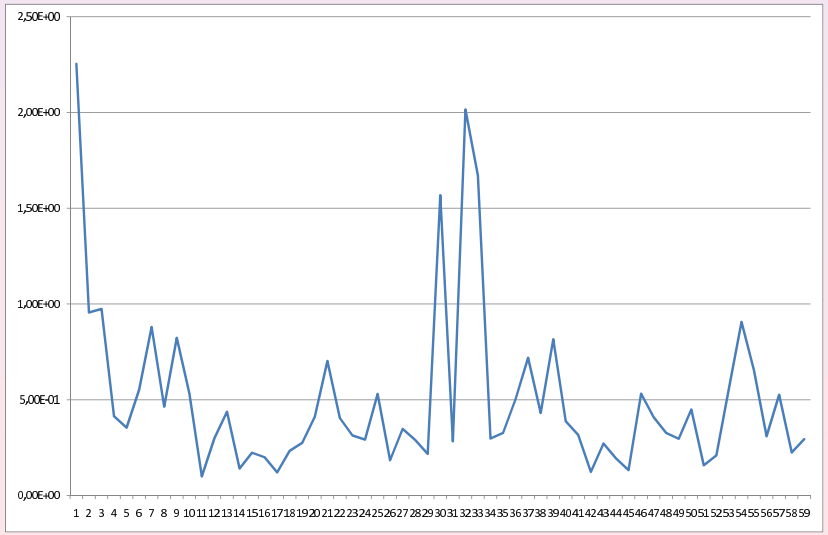


Data04: FFT=5, prebleach=4: y-schema, y-solution, p, J

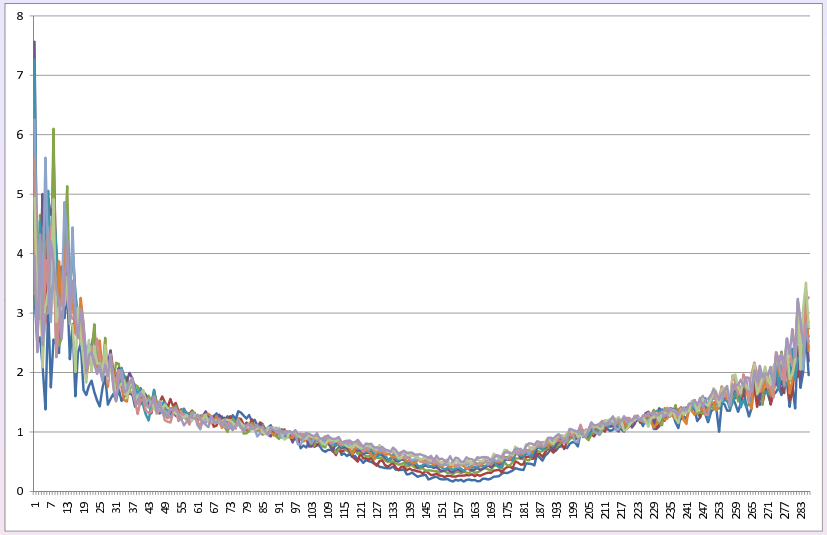


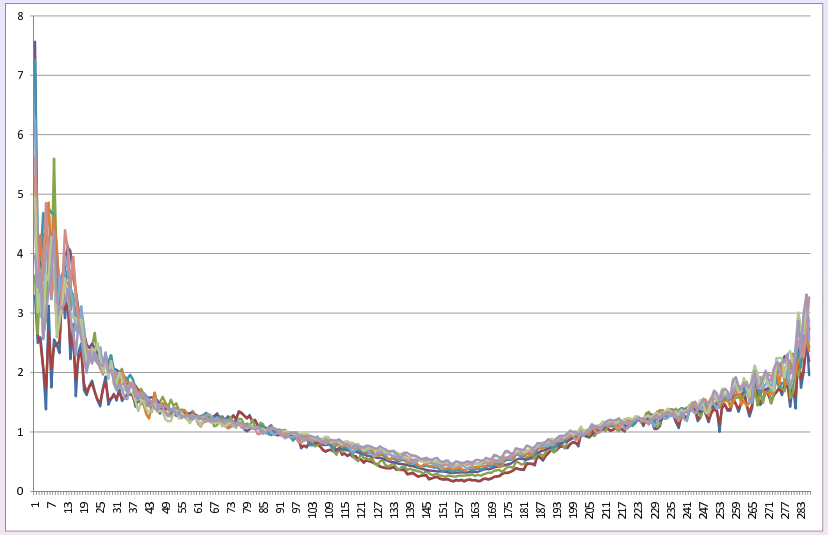


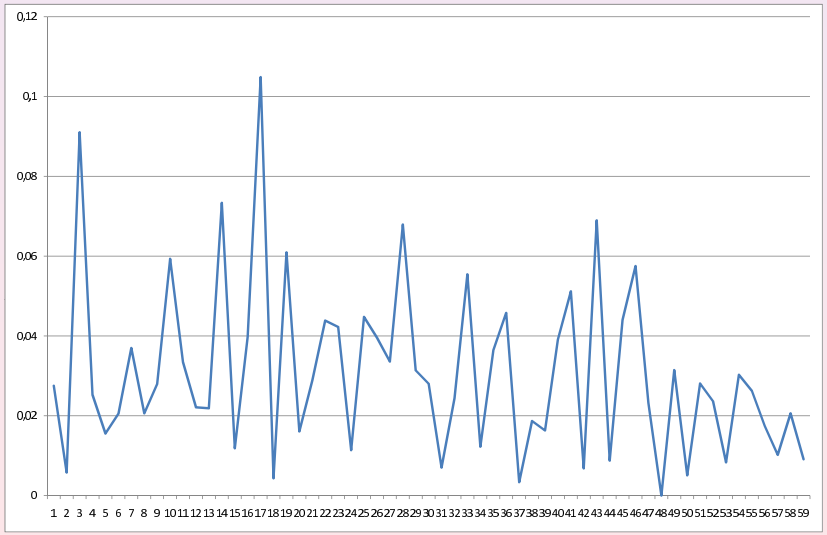


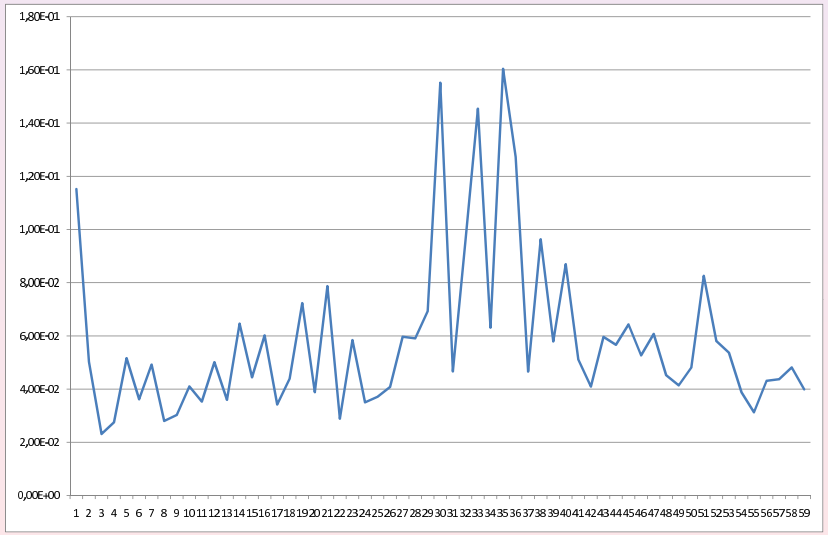


Data05: FFT=0, prebleach=-4: y-schema, y-solution, p, J

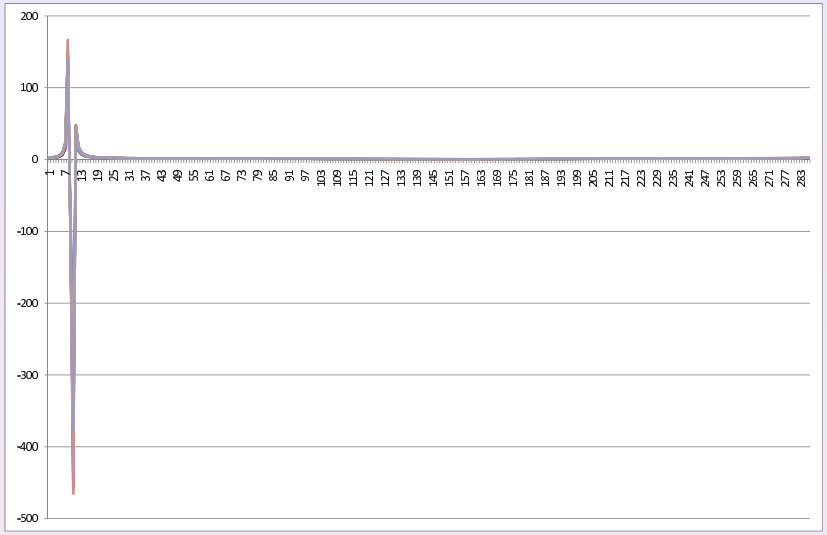


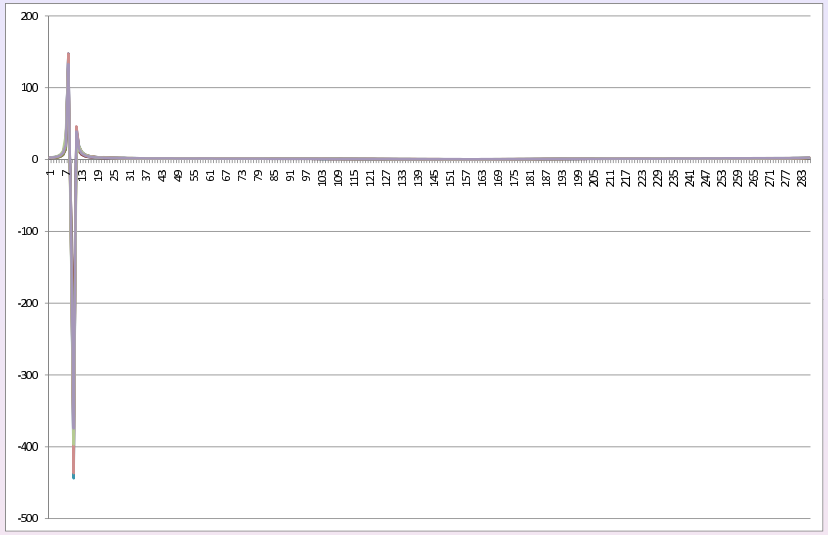


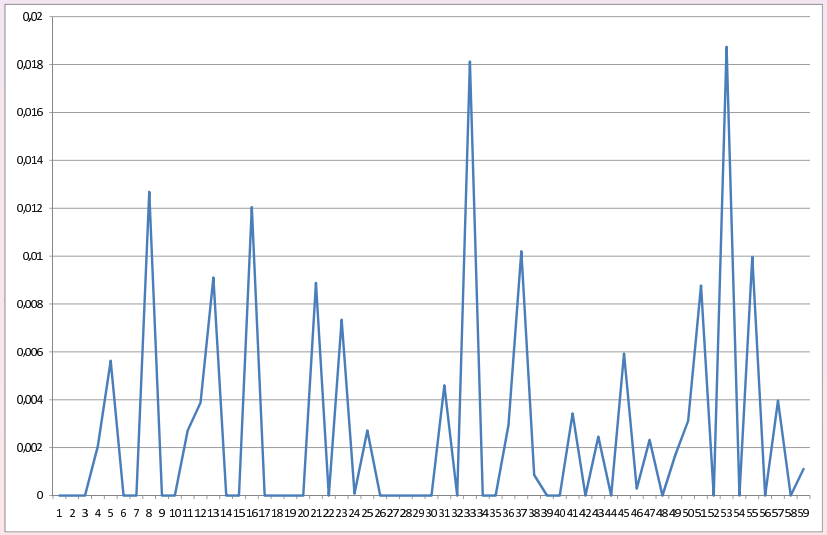


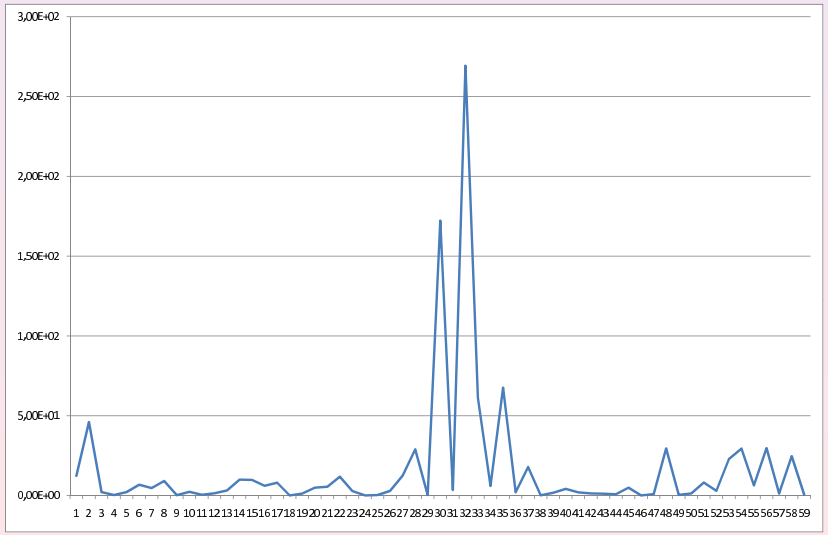


Data06: FFT=5, prebleach=-4: y-schema, y-solution, p, J

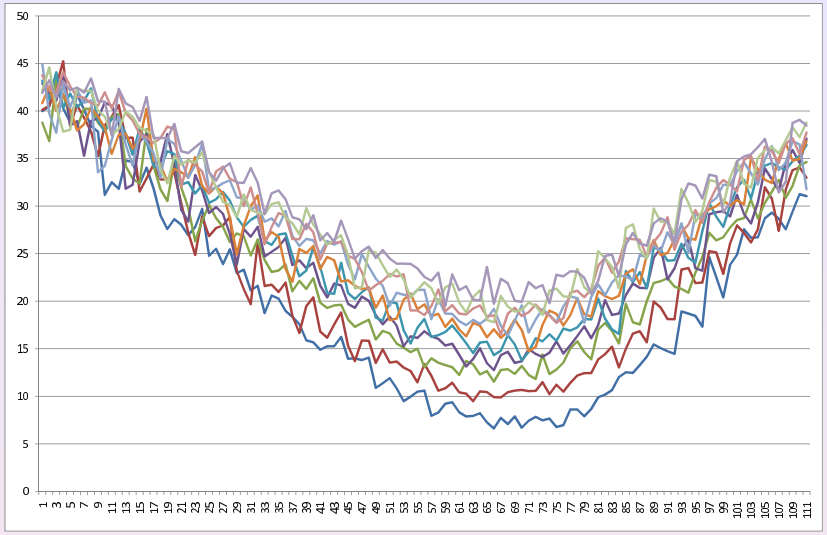


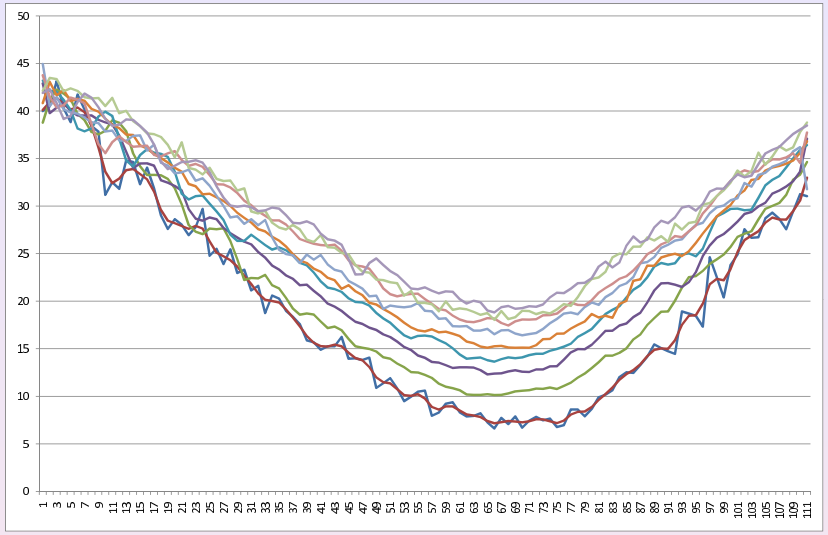


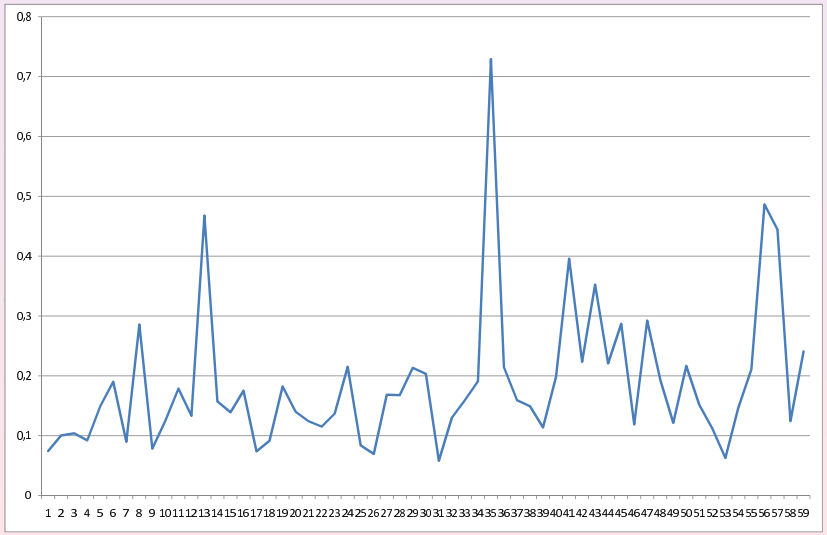


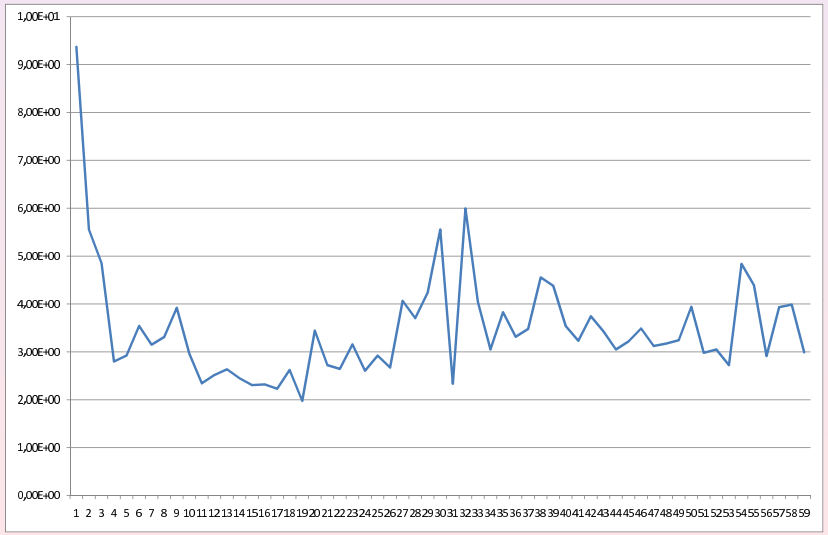


Data07: FFT=0, prebleach=0: y-schema, y-solution, p, J

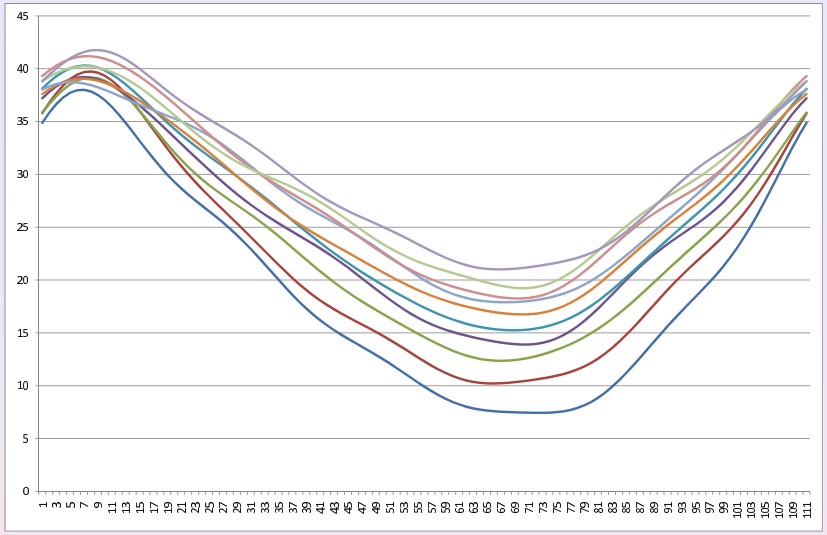


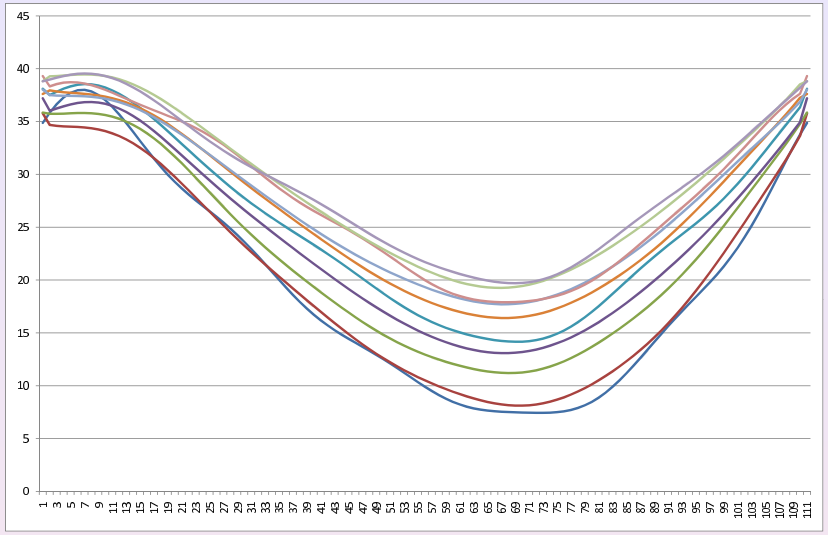


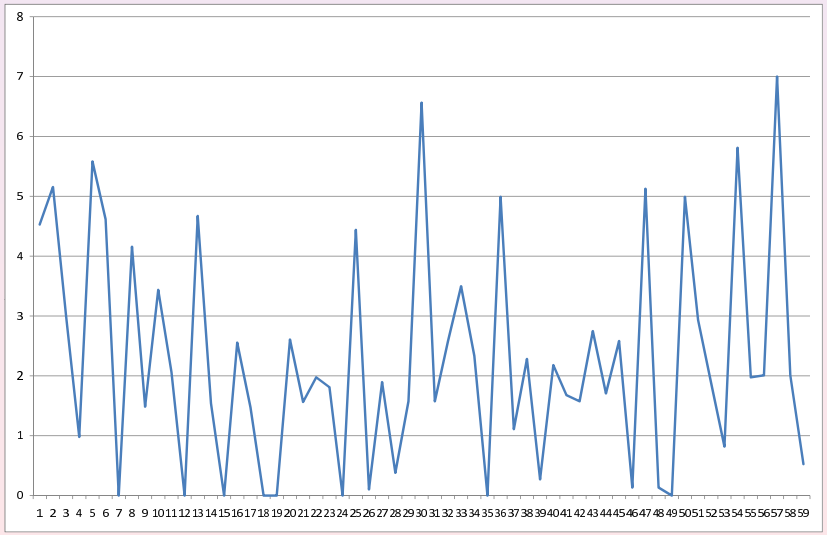


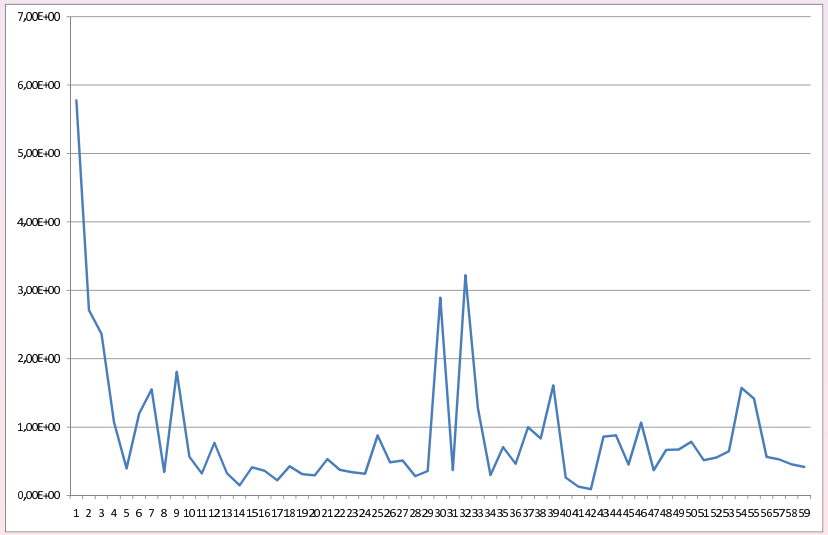


Data08: FFT=5, prebleach=0: y-schema, y-solution, p, J

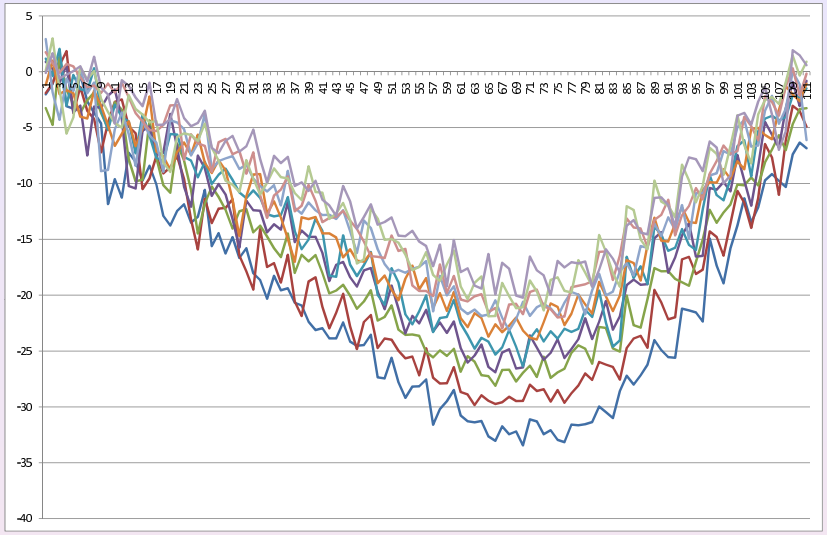


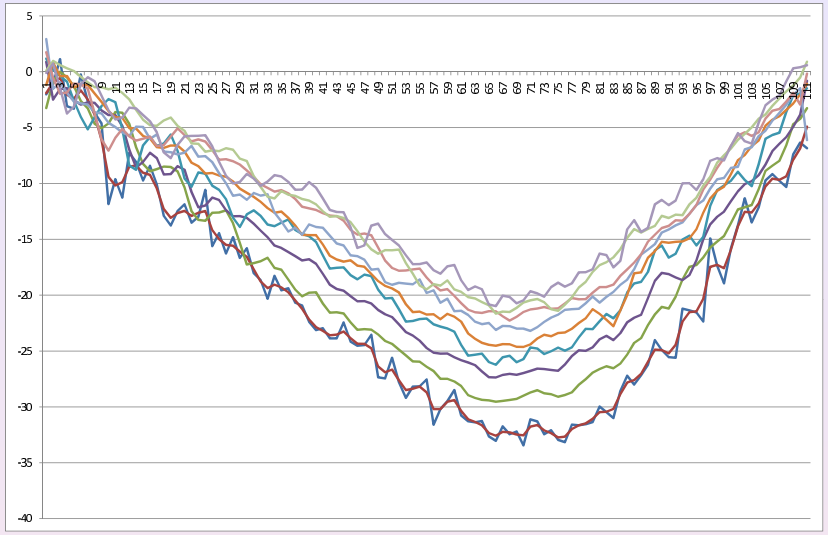


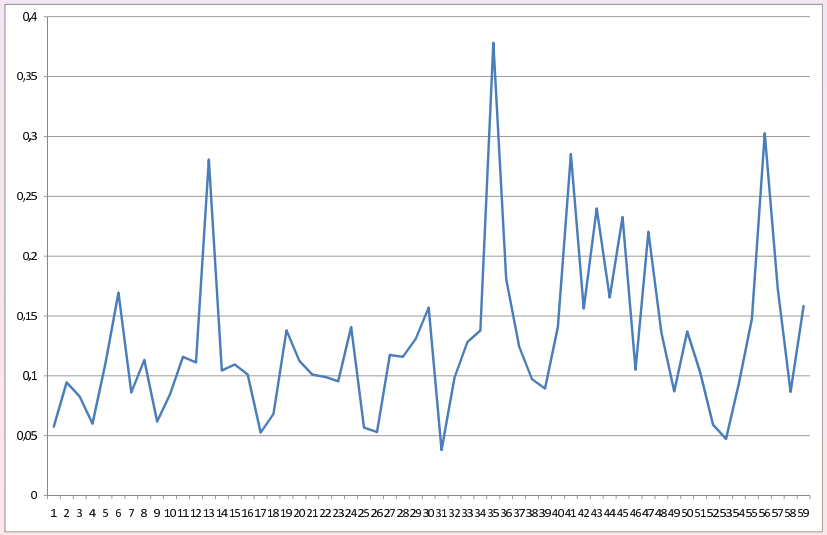


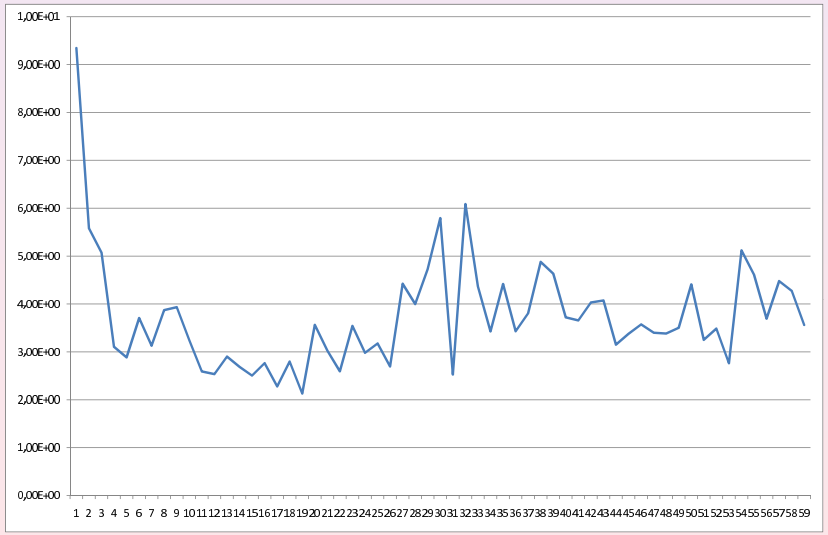


Data09: FFT=0, prebleach=4: y-schema, y-solution, p, J

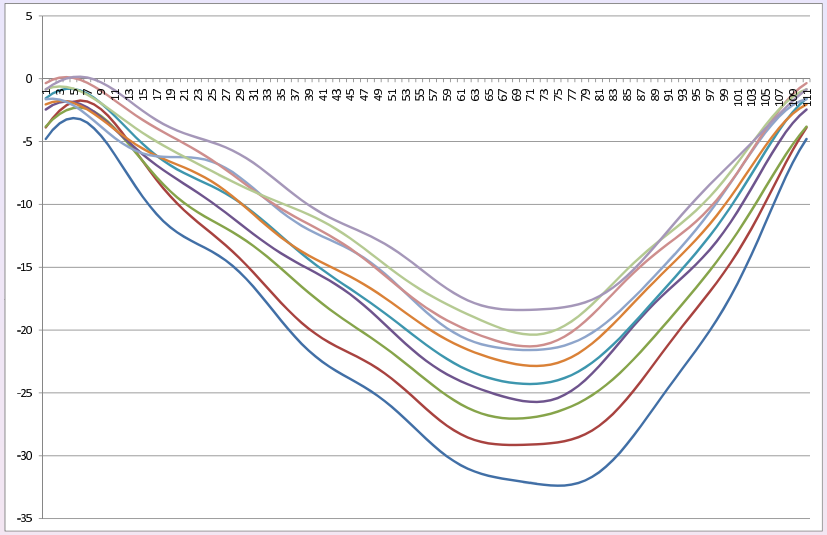


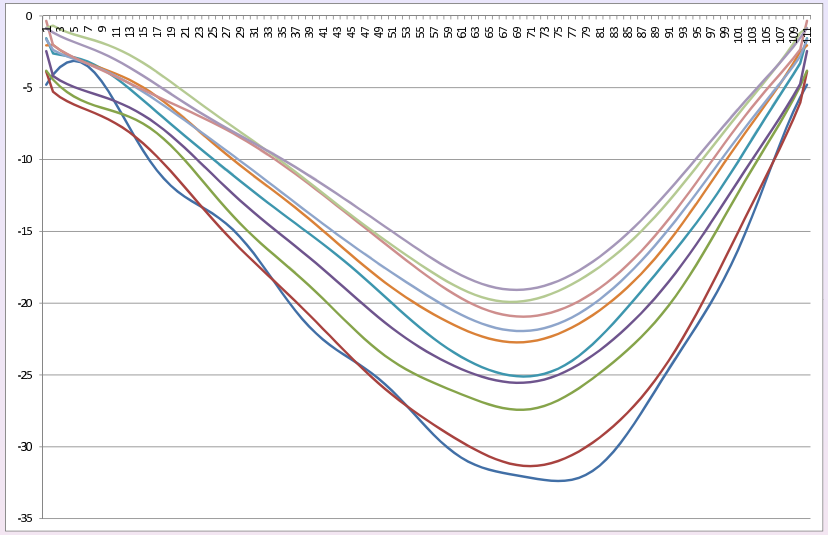


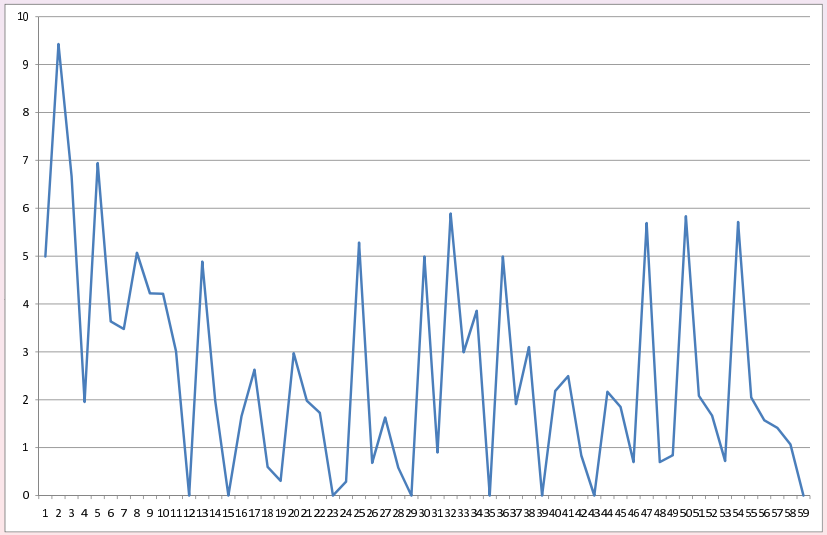


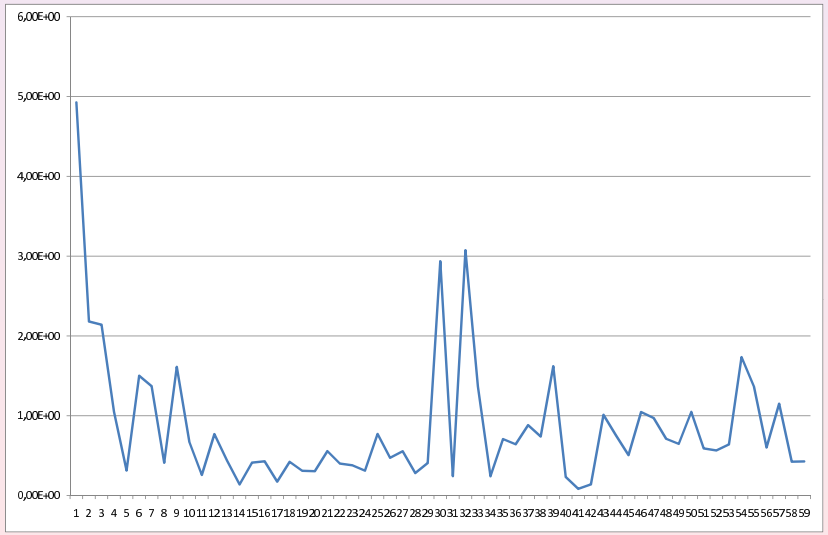


Data10: FFT=5, prebleach=4: y-schema, y-solution, p, J

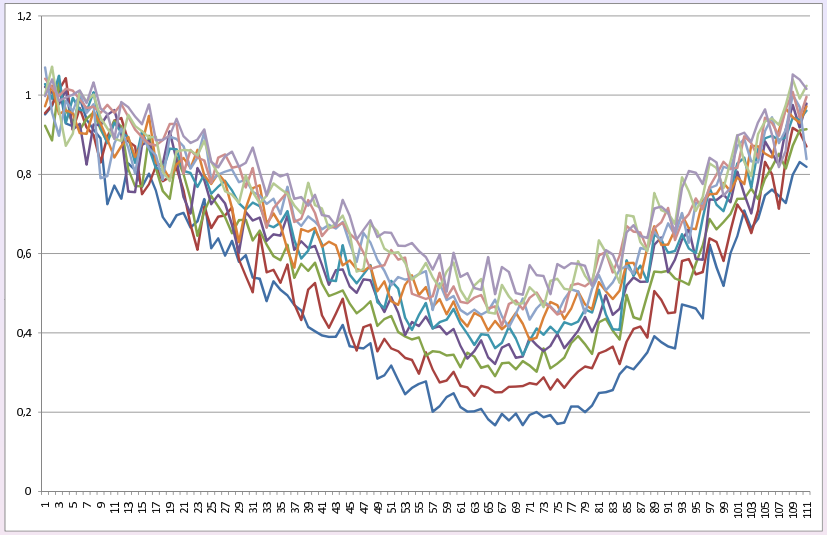


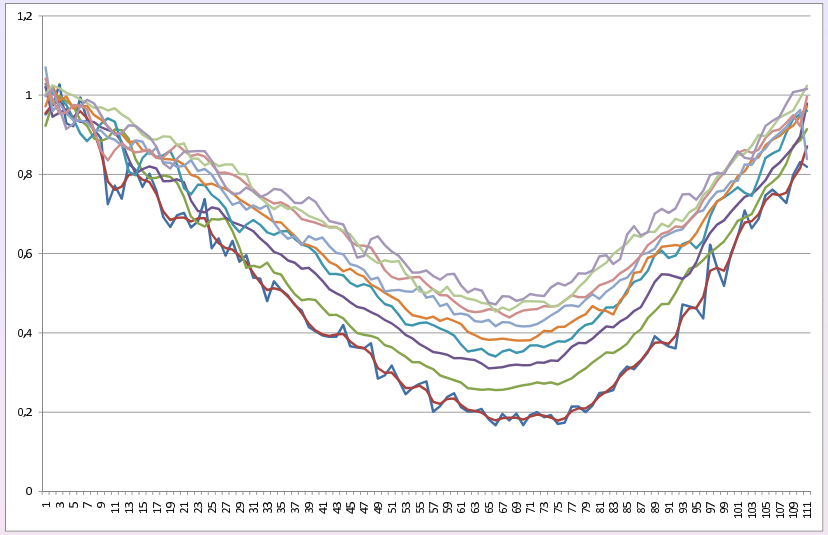


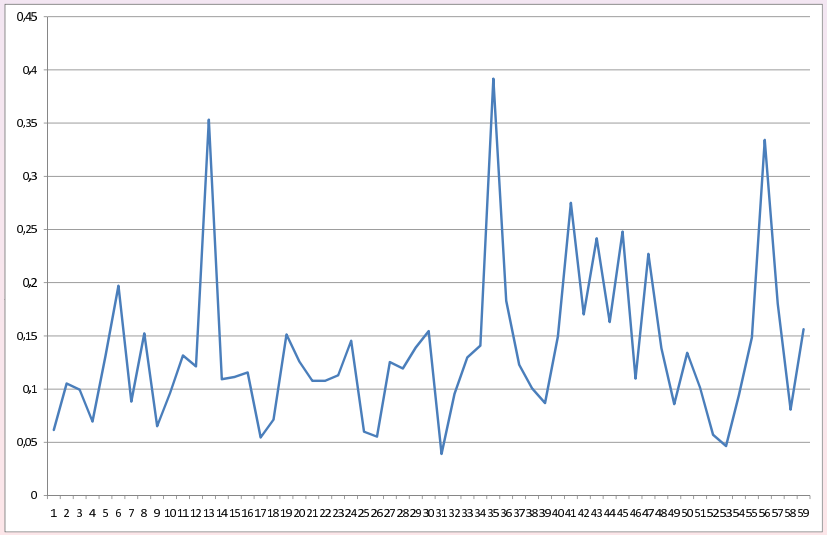


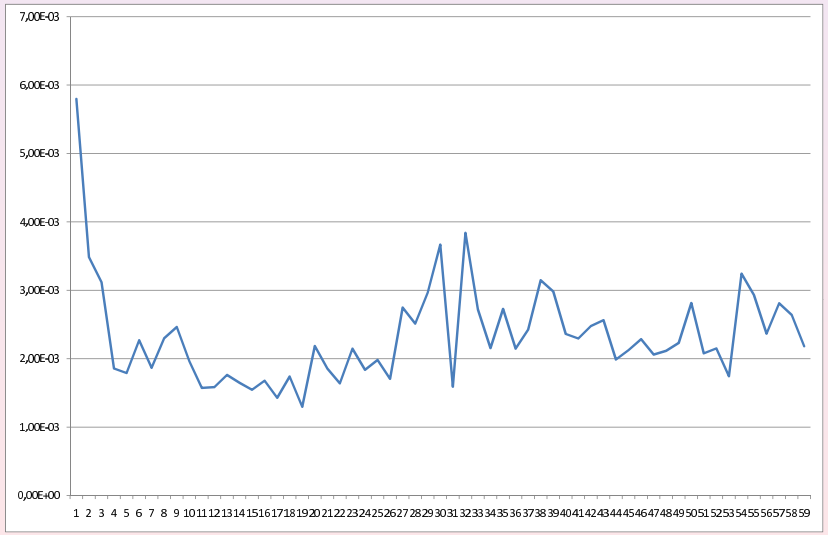


Data11: FFT=0, prebleach=-4: y-schema, y-solution, p, J

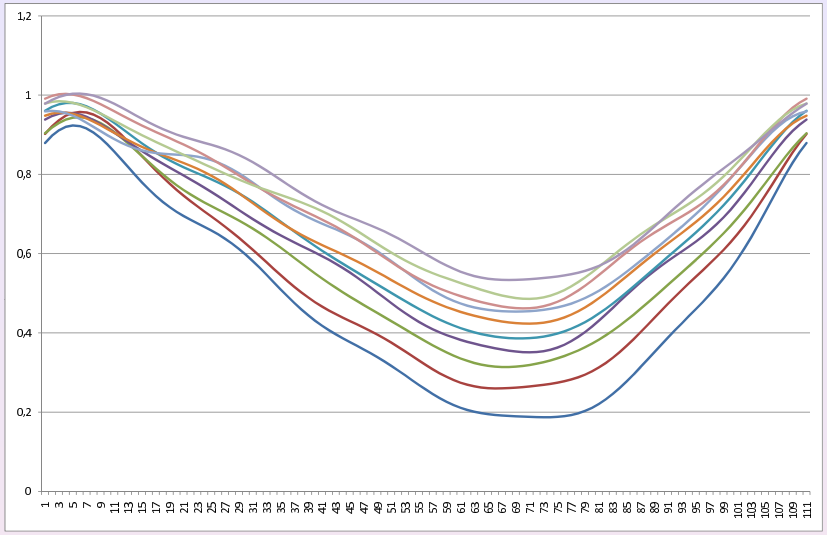


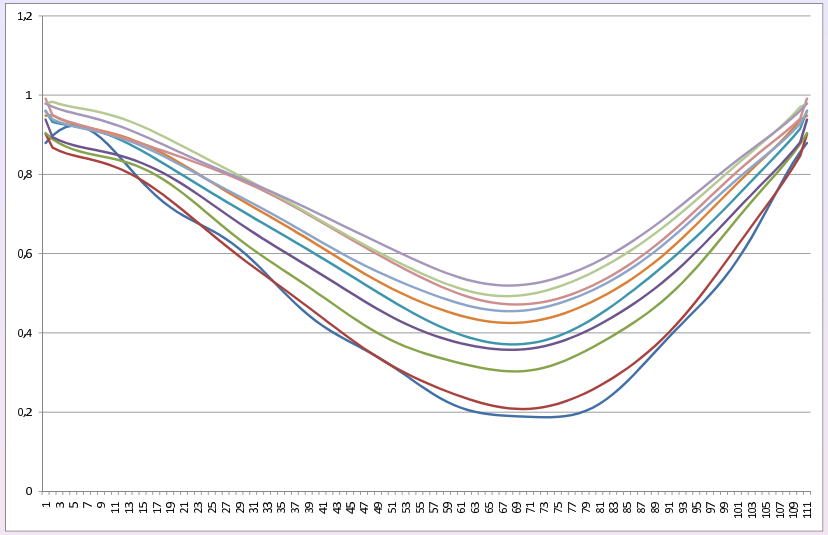


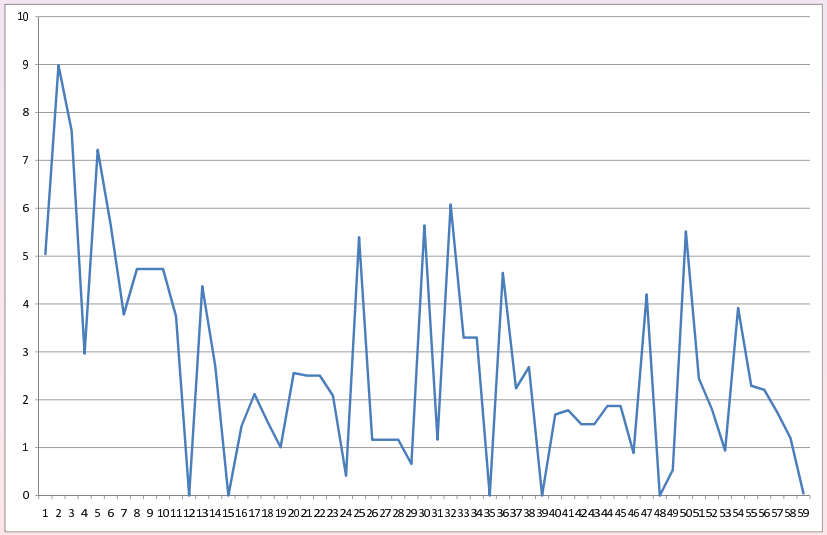


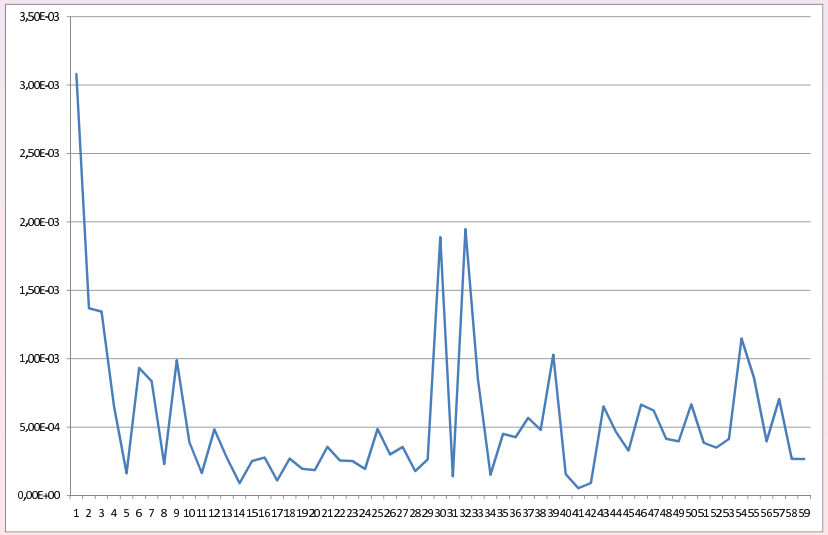


Data12: FFT=5, prebleach=-4: y-schema, y-solution, p, J

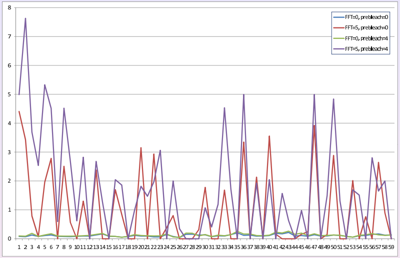


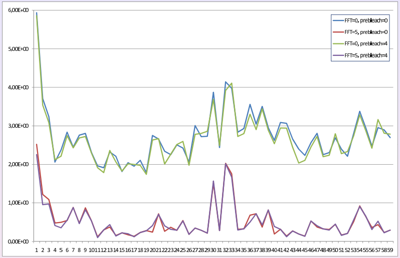


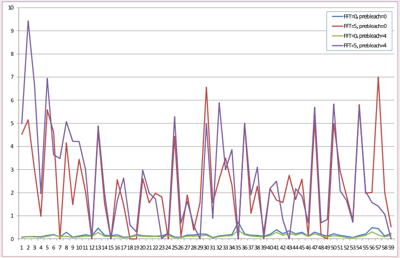


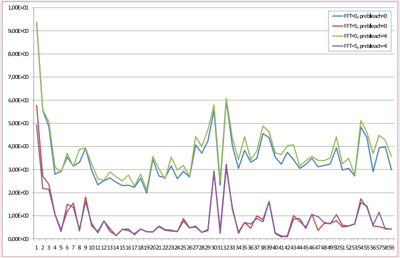


Data: A=1, B=286: p, J; A=90, B=200: p, J









- Our program is actually under testing, however, for the previously known diffusion coefficient and the data simulated by the random walk model it computes correct results. Afterward, we determined the diffusivities for the real data of FRAP measurements (with the red algae *Porphyridium cruentum*). The range of result $10^{-14}\mathrm{m}^2\mathrm{s}^{-1}$ is in agreement with reference values.
- Our method for phycobilisomes diffusivity estimation from FRAP data improves on other models by accounting for experimentally measured post-bleaching fluorescence profiles and time-dependent boundary conditions, and can include also a reaction term to account for the low level bleaching during scanning and the time varying fluorescence signal.
- Although the minimization of function J is a one-dimensional optimization problem, finding an optimal solution p is quite a difficult task (J seems to be flat near the solution). In the near future, we would like to improve our method by an adequate assessment of the measurement noise and by an implementation of a more robust optimization procedure.