# ARTIFICIAL NEURAL NETWORK AS A TOOL FOR PRELIMINARY ANALYSIS OF TIME RESOLVED FLUORESCENCE DATA

**Petr V. Nazarov, Marina V. Repich, Vladimir V. Apanasovich**

*Department of Systems Analysis,*
*Faculty of Radiophysics and Electronics,*
***Belarusian State University***
*Nezalezhnasti Ave 4, 220050,*
*Minsk, Belarus.*
*Tel/fax: +375 17 2789345*
*E-mail: nazarov@bsu.by*

# Time-resolved Fluorescence Spectroscopy

## Introduction

The **time resolved fluorescence spectroscopy** is a very important experimental tool to study the complex biomolecular objects and systems, including liquid crystals, membranes, proteins, DNA, etc. This method provide detailed information about the structure and dynamic of those systems [1]. The analysis of spectroscopic data may be complex because of several reasons: there could be a number of unknown parameters in an experimental system; almost all dependencies between them and the processes taking place are non-linear; fluorescence data are distorted by noises and inaccuracies of a registration system. These facts impel to analyze experimental data via the multi-parametric optimization approach (fitting) [2].

For the successful application of the fitting procedure, the **selected model** of fluorescence must be an **adequate** and the **initial estimation** for its parameters should be **sufficiently good**. Consequently, the tasks of model recognition and initial estimations arise. This problems belong to the **"inverse problems"** and therefore their solution is unstable and very much noise sensitive.

**Artificial neural networks (ANN)** are well-known for their robustness, noise stability and ability to approximate any smooth function. Therefore it was decided to test their applicability to the problems of model selection and initial estimations.

## Goal

To develop a noise-stable neural network based algorithms for the model selection (recognition) and preliminary fluorescence data analysis (parameter extraction).

## Objectives

- Development of the preprocessing scheme for reduction of time-resolved fluorescence data
- Selection of the ANN structure
- Demonstration the possibility of the fluorescence model recognition by ANN
- Demonstration the possibility of the initial estimation by ANN

# Problem Definition

The fluorescence decay of a molecular ensemble depends on lifetimes of fluorescent molecules and interactions between them. For the simplest case of non-interacting mixture of fluorophores with different lifetimes the decay curve can be expressed as a **sum of exponents** (Eq 1). In the case of Förster type of energy transfer between molecules the fluorescence can be presented via **stretched exponents** (Eq 3). And in the most general case the behavior become rather complex (Eq.3), where *G* **cannot be expressed analytically** in some situations.

$$f(t) = I_0 \sum_j e^{-k_j t} \quad (1) \qquad f(t) = I_0 e^{-ct^n} \quad (2) \qquad f(t) = I_0 e^{-G(t,k_r,\ldots)} \quad (3)$$

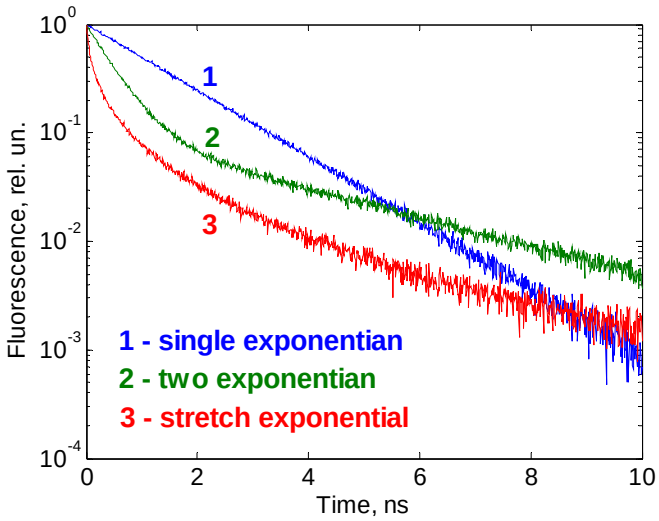



*Fig.3. Types of fluorescence decays*

The original fluorescence decay *f*(*t*) caused by the system itself is distorted by the **convolution** with the excitation pulse together with different detector time lags (fig.4).

The extraction of the original *f*(*t*) by deconvolution is noise unstable operation and should be avoided. Unfortunately standard approaches to preliminary analysis (i.e. Laplace transform) work on deconvoluted data.

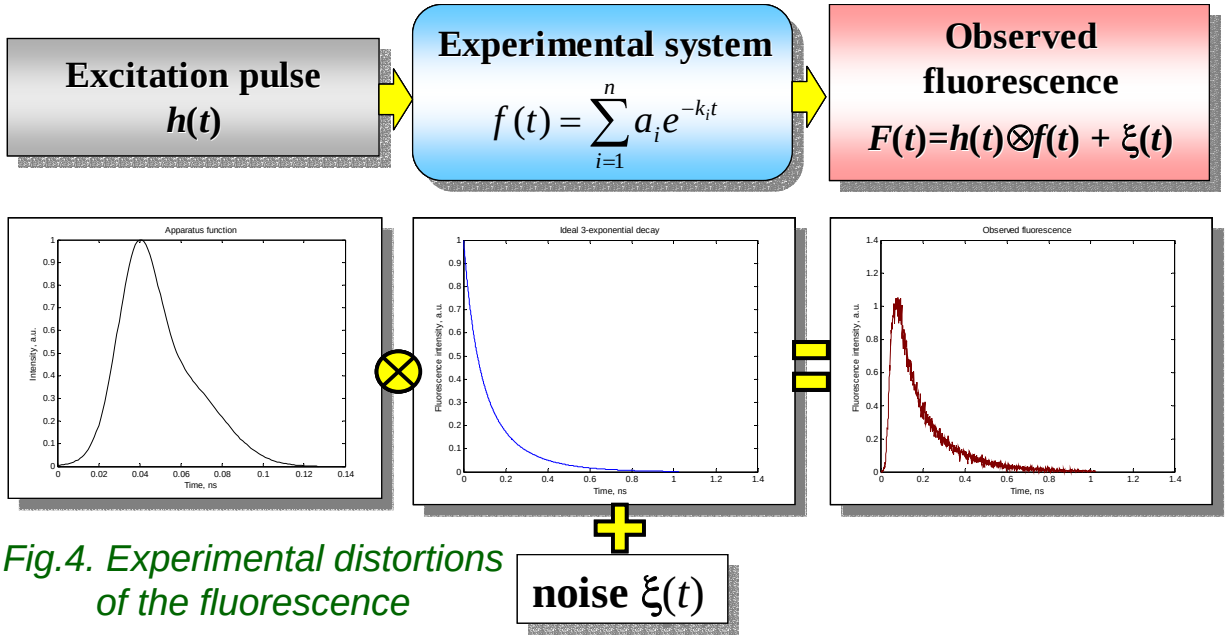



*Fig.4. Experimental distortions of the fluorescence*

# Approaches to Fluorescence Data Analysis

The general approach to fluorescence data analysis is presented in fig. 1.

Let experiments to be performed on a system (1) which has some unknown parameters $P_x$. The result is a set of experimental data (2). Using *a priory* information about the system and experimental data, the model (type of the decay) is assumed (3). By this model the initial estimations of $P_x$ should be made (4). After that the precise parameters determination can be performed using, for instance, simulation-based fitting technique [2].

The simulation-based fitting approach is illustrated in fig. 2. Again the experiment on the system (1) provides some experimental data (2). Then estimation of $P_x$ (3) is used in the model (4) to produce simulated data (5). The fitting algorithm (6) modifies estimation of $P_x$ to minimize differences between simulated and experimental data. The process iteratively continues till the stop criterion is reached.

Obviously, the model selection and initial estimation are the key points in fluorescence data analysis.

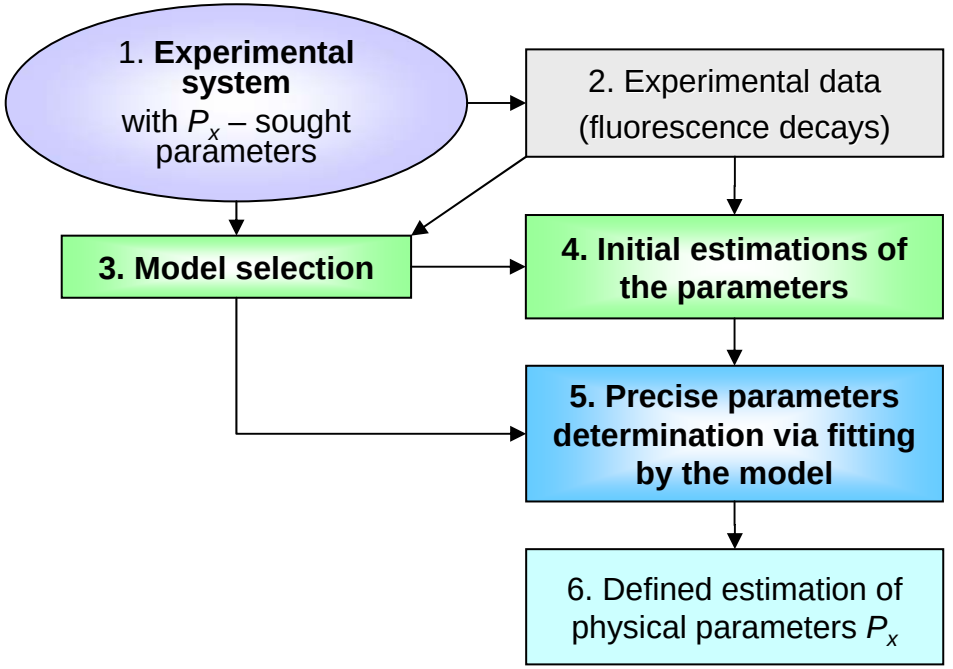



*Fig.1. General scheme of fluorescence data analysis*

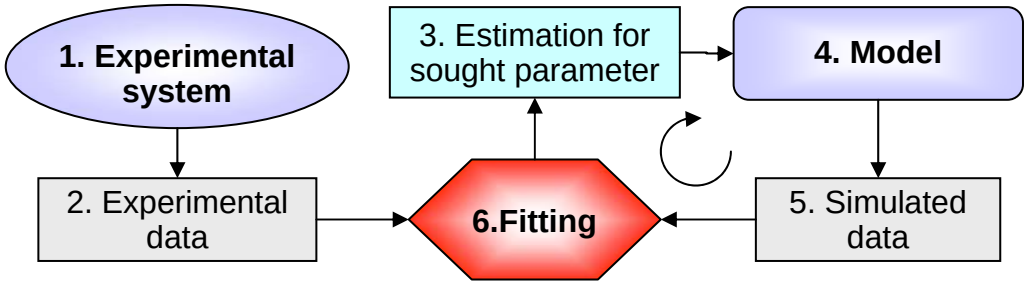



*Fig.2. Precise parameters determination via fitting*

# Artificial Neural Network Approach

Artificial neural networks (ANN) are widely used for the analysis of distorted data. The key features of ANN are their noise stability, ability to be trained directly on experimental data (with all their distortions) and generalization [3].

The general scheme of ANN approach to the inverse problem solution is showed in fig.5. Experimental system (2) can be considered as a “black box” transferring physical parameters (1) into experimental data (3). Then the data are utilized by an ANN to estimate the original values of sought parameters (4).

ANN can be used in the same way to recognize the model of the fluorescence decay (exponential, stretched exponential, etc.)

Before analysis by ANN the decays should be preprocessed to reduce the dimensionality of the data. For example this can be performed in the following way. Let each decay contain 1024 channels. Data can be reduced to 8 points by averaging inside intervals containing 8, 8, 16, 32, 64, 128, 256, 512 channels (fig.6).

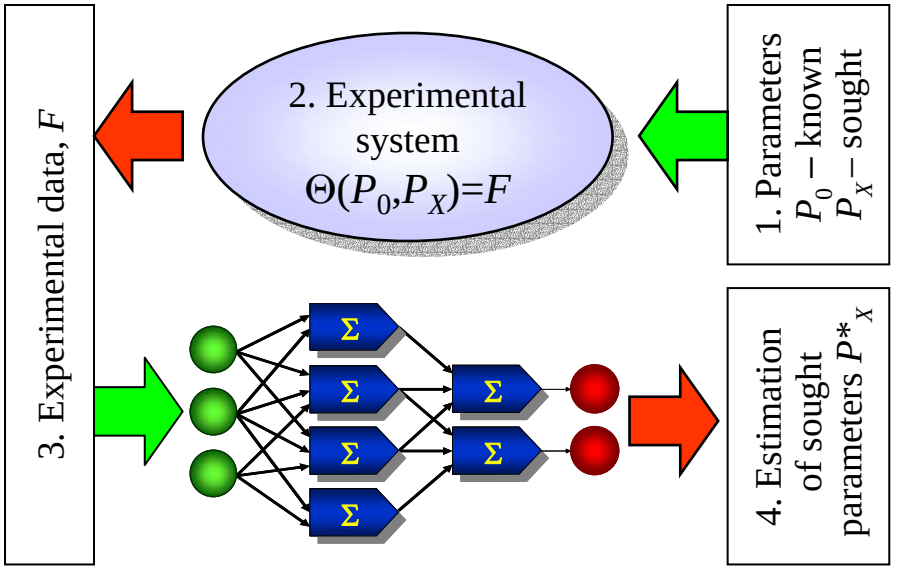



*Fig.5. Scheme of ANN application for data analysis*

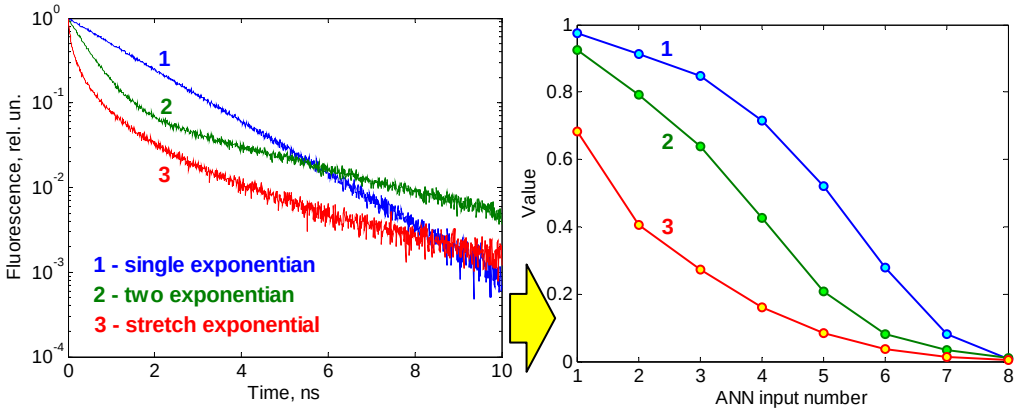



*Fig.6. Preprocessing and data reduction*

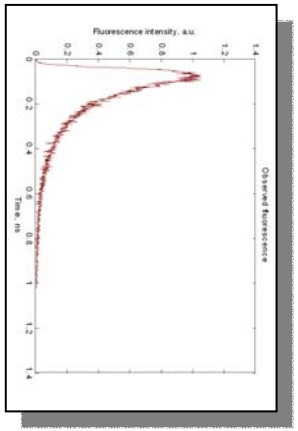
# Model Selection

## Numerical Experiment

**Data:** Three models were considered: 1-, 2- and 3-exponential. The data were simulated, convoluted with a smooth pulse function and distorted by noise (fig. 6). The training set consisted of 1000 training pairs.

**ANN structure:** Three layer feed-forward ANN (perceptron) with 8 neurons in hidden layers and three outputs. Preprocessed decays were given to the inputs. The value in the 1st, 2nd and 3rd outputs provided the probability of the 1-, 2- and 3-exponential models. The maximal probability defined the final decision about the model.

**Training:** ANN was trained using back-propagation error method with Levenberg-Marquartd modification in Matlab.



ANN classification

"probability" of 1 exponent model

"probability" of 2 exponent model

"probability" of 3 exponent model

*Fig.7. Model recognition*

## Results and Discussion

The approach was tested on 1000 decays. The results are illustrated in fig. 8.

The single exponential model was determined correctly with the probability of 98%, for two-exponential model – 94% and for three-exponential model – 92%.

The high error for three-exponential model can be explained by the fact that its fluorescence rate constants can have similar values and therefore be recognized as 1- and 2- exponential models.

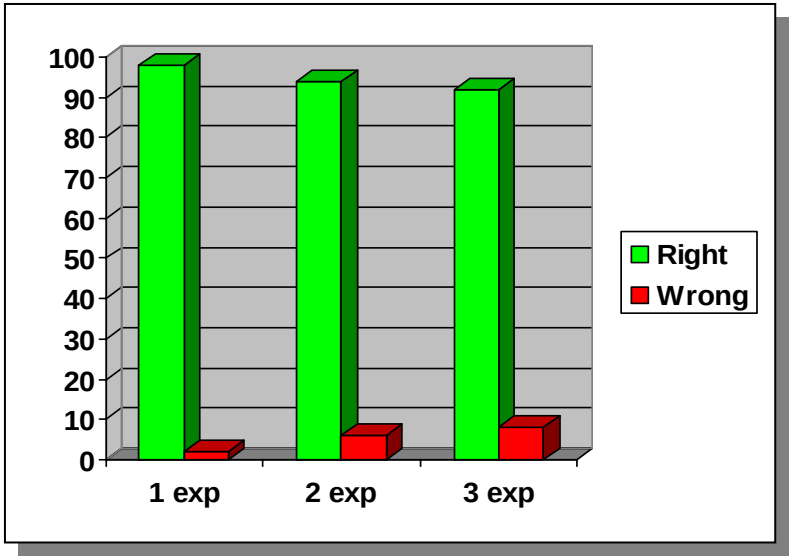



*Fig.8. Results of the model recognition*

# Initial Estimations

## Numerical Experiment

**Data:** Three exponential decays were considered with rate constants $k_1<k_2<k_3$. The data were simulated, convoluted with a smooth pulse function and distorted by noise (fig. 6). The training set consisted of 2000 training pairs.

**ANN structure:** Three layer feed-forward ANN (perceptron) with 16 neurons in hidden layers and three outputs. Preprocessed decays were given to the inputs. The value in the 1st, 2nd and 3rd outputs provided the normalized estimation of $k_1$, $k_2$, $k_3$.

**Training:** ANN was trained using back-propagation error method with Levenberg-Marquartd modification in Matlab.

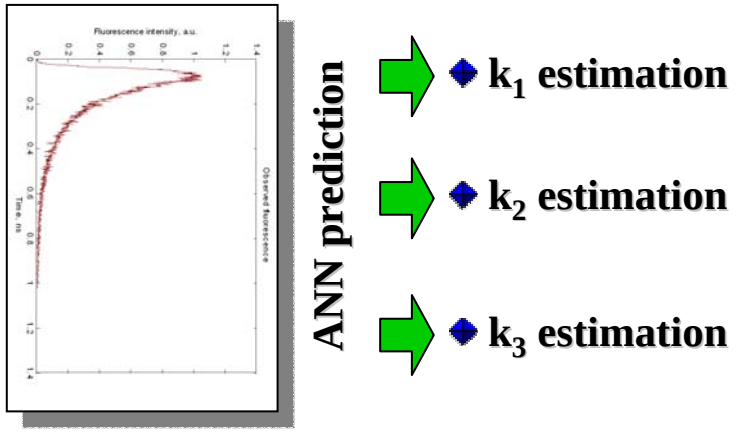



*Fig.9. Preliminary parameter estimation*

## Results and Discussion

The approach was tested on 1000 decays. The results are illustrated in fig. 8.

Three error regions were considered: <10%, 10-20% and 20-30%. For more then 960 decays the errors in defined rate constants were below 20%, and for 700 decays it was even below 10%.

The better prediction was obtained for the first exponential component with longer lifetime (smaller rate).

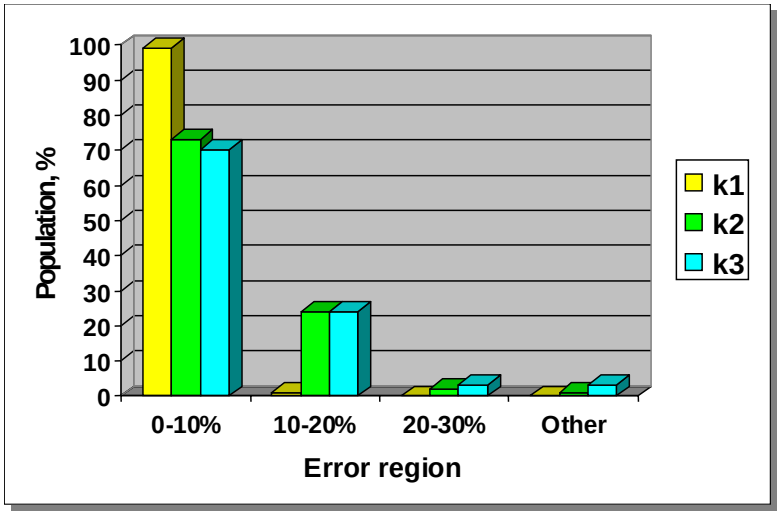



*Fig.10. Errors distribution for three rate constants after analysis of noisy and convoluted data*

# Conclusions

- Artificial neural networks are able to recognize the model of the fluorescence decay with high precision (average accuracy >95%)
- Artificial neural networks can be used to produce initial estimations of the sought parameters hidden in time-resolved fluorescence data. Average precision of the definition of rate constants with <10% error is ~81%
- ANN can be applied to analyze complex (stretch exponential, etc) decays, in the convoluted form. The proposed technique is a noise stable one.

## Literature


1. Lakowicz JR. 1999. *Principles of fluorescence spectroscopy*. New York: Kluwer Academic/Plenum Publishers.
2. Nazarov PV, Apanasovich VV, Lutkovski VM, Yatskou MM, Koehorst RBM, Hemminga MA. 2004. Artificial neural network modification of simulation-based fitting: application to a protein-lipid system. *J. Chem. Inf. Comput. Sci.* 44:568-574.
3. Wasserman, PD. 1989. *Neural Computing Theory and Practice*. New York: Van Nostrand Reinhold