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

R. Likhonina, E. Suzdaleva, I. Nagy

Comparison of mixture-based classification with the data-dependent pointer model for various types of components

2355 August 2016

Project GAČR GA15-03564S

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

Contents

1 Introduction

1.1 Motivation and state of the art

The presented research report deals with a mixture-based classification problem, which is one of domains of unsupervised model-based classification methods. Often they are based on clustering solutions applied to the untrained data set, see e.g., [1, 2, 3], etc.

The focus of the report is on classification using the mixture models, which consist of components in the form of probability density functions (pdfs) describing individual regimes of working a considered system and a model of their switching. The mixture-based classification starts from some pre-specified (mostly resulted from the initial data analysis) locations of components and performs a search for density clusters in the data space with the aim of fitting component models to data.

The key part of the mixture-based classification in this context is determining the component which describes the active regime of the system. This brings a series of involved subtasks, such as estimation of parameters of components and of the switching model. Thus the mixture estimation should be applied. There exists a variety of various approaches in this field, namely: (i) estimation based on the use of the EM algorithm [4, 5, 6]; (ii) Variational Bayes methods [7, 8]; (iii) Markov Chain Monte Carlo (MCMC) methods [10, 11, 12]; and (iv) recursive Bayesian estimation algorithms [15, 13, 14, 16, 17, 18], etc. Here, the last part of the enumerated approaches is used, which is directed at algebraic computing the statistics of the involved component distributions avoiding applying the numerical techniques.

1.2 Aim of experiments

The research report describes an experimental comparison of the classification quality for a series of types of mixture components, namely:

- categorical components;
- uniform components;
- exponential components;
- state-space components with known parameters
- and normal components.

The dynamic and static configuration of the data-dependent switching model is used. For the first of them all types of components are considered. The static switching model is taken with normal components.

The aim of the comparison is to explore advantages of using the data-dependent switching model from the point of view of classification.

2 Estimation scheme

This section introduces the considered models and schematically summarize the mixture estimation algorithm used for the classification task.

2.1 Models

Let's consider a multi-modal system, which at each discrete time instant $t = 1, 2, ...$ generates continuous data y_t and/or discrete data z_t with the set of its possible values $\{1, 2, \ldots, m_z\}$. It is assumed that the observed system works in m_c working modes indicated by values of the unmeasured dynamic discrete variable $c_t \in \{1, 2, \ldots, m_c\}$, which is called the pointer [15], and each of the pointer values also depends on values of the measured variable z_t. The variable y_t is in general the N_y-dimensional column vector.

The observed system is described by a mixture model consisting of m_c components in the form of the following pdfs

$$f\left(\textbf{the modeled variable } |\Theta, c_t = i\right),\ i \in \{1, 2, \ldots, m_c\}, \tag{1}$$

where $\Theta = \{\Theta_i\}_{i=1}^{m_c}$ is a collection of unknown parameters of all components, and Θ_i includes parameters of the i-th component according to its distribution (see below) in the sense that $f\left(y_t|\Theta, c_t = i\right) = f\left(y_t|\Theta_i\right)$ for $c_t = i$.

The component, which describes data generated by the system at the time instant t is said to be active. A distribution of the component (1) is chosen according to the nature of measurements and made assumption about modeled variables. Specifically, here the pdf (1) is taken in the following forms.

2.1.1 Components

The categorical components are considered for the case of discrete measurements $z_t \in \{1, 2, \ldots, m_z\}$. Each component (1) is taken as

$$f(z_t = l | \psi^z_{t-1} = q, \beta, c_t = i) \equiv \tag{2}$$

	$z_t = 1$	$z_t = 2$	$\cdots$	$z_t = m_z$
$\psi^z_{t-1} = 1$	$(\beta_{1\|1})_i$	$(\beta_{2\|1})_i$	$\cdots$	$(\beta_{m_z\|1})_i$
$\psi^z_{t-1} = 2$	$(\beta_{1\|2})_i$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\psi^z_{t-1} = m_\psi$	$(\beta_{1\|m_\psi})_i$	$\cdots$	$\cdots$	$(\beta_{m_z\|m_\psi})_i$

with the regression vector $\psi^z_{t-1} = [z_{t-1}, \ldots, z_{t-n_z}]'$ of the length n_z, and with the parameter β_i corresponding to the component labeled by the value $c_t = i$, such that $\{\beta_i\}_{i=1}^{m_c} \equiv \beta$. Denotation $q \in \{1, 2, \ldots, m_\psi\}$ belongs to a configuration of the used regression vector. Here β_i is a matrix of non-negative probabilities $(\beta_{l|q})_i$ of the value $z_t = l$ conditioned by the regression vector configuration $\psi^z_{t-1} = q$ for $c_t = i$.

The pdf (1) with the uniform distribution under assumption of the independence of entries of the vector y_t is taken $\forall i \in \{1, 2, \ldots, m_c\}$ as follows:

$$f\left(y_t | L, R, c_t = i\right) = \begin{cases} \frac{1}{R_i - L_i} & \text{for } y_t \in (L_i, R_i), \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

i.e., here $\{L_i, R_i\} \equiv \Theta_i$, and their entries $(L_l)_i$ and $(R_l)_i$ are respectively minimal and maximal bounds of the l-th entry $y_{l;t}$ of the vector y_t for the i-th uniform component, and $l = \{1, 2 \ldots, N_y\}$.

The exponential distribution of components (1) is specified as

$$\left(\prod_{l=1}^{N} (a_l)_i\right) \exp\left\{-a_i'(y_t - b_i)\right\}, \tag{4}$$

i.e., here $\{a_i, b_i\} \equiv \Theta_i$, and $(a_l)_i > 0$ and $(b_l)_i \in R$ are the l-th entries of the N_y-dimensional vectors a_i and b_i respectively with $l = \{1, 2 \ldots, N_y\}$. Currently the independence of entries of the vector y_t is assumed.

The state-space components (1) are presented as the following pdfs:

$$f(x_t | x_{t-1}, u_t, c_t = i), \quad f(y_t | x_t, u_t, c_t = i), \quad \forall i \in \{1, 2, \ldots, m_c\}, \tag{5}$$

where x_t is the unmeasurable state to be estimated, and (5) are linear models existing for each value of the categorical state c_t. is written as the following two equations:

$$x_t = M x_{t-1} + N u_t + \omega_t, \tag{6}$$

$$y_t = A x_t + B u_t + v_t, \tag{7}$$

where M, N, A and B are matrices of parameters of appropriate dimensions supposed to be known; ω_t and v_t are the process and the measurement Gaussian white noises with the zero means and the covariance matrices R_ω and R_v respectively, which are supposed to be known.

The normal components are presented as the following pdfs:

$$(2\pi)^{-N_y/2} |r_i|^{-1/2} \exp\left\{-\frac{1}{2}[y_t - \theta_i]' r_i^{-1} [y_t - \theta_i]\right\}, \tag{8}$$

where θ_i is the mean vector of the i-th normal component and r_i is the covariance matrix.

2.1.2 Dynamic Data-dependent Pointer Model

Switching the active components (1) can be described by the dynamic data-dependent model of the pointer

$$f\left(c_t = i|\alpha, c_{t-1} = j, z_t = k\right) = \tag{9}$$

	$c_t = 1$	$c_t = 2$	$\cdots$	$c_t = m_c$
$c_{t-1} = 1$	$(\alpha_{1\|1})_k$	$(\alpha_{2\|1})_k$	$\cdots$	$(\alpha_{m_c\|1})_k$
$c_{t-1} = 2$	$(\alpha_{1\|2})_k$		$\cdots$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$c_{t-1} = m_c$	$(\alpha_{1\|m_c})_k$		$\cdots$	$(\alpha_{m_c\|m_c})_k$

where the unknown parameter α is the $(m_c \times m_c)$-dimensional matrix, which exists for each value $k \in \{1, 2, \ldots, m_z\}$ of z_t. Its entries $(\alpha_{i|j})_k$ are non-negative probabilities of the pointer $c_t = i$ under condition that the previous pointer $c_{t-1} = j$ with $i, j \in \{1, 2, \ldots, m_c\}$ and $z_t = k$.

2.1.3 Static Data-dependent Pointer Model

Switching the components can be also described by the static data-dependent model of the pointer

$$f\left(c_t = i|\alpha, z_t = k\right) = \tag{10}$$

	$c_t = 1$	$c_t = 2$	$\cdots$	$c_t = m_c$
$z_t = 1$	$\alpha_{1\|1}$	$\alpha_{2\|1}$	$\cdots$	$\alpha_{m_c\|1}$
$z_t = 2$	$\alpha_{1\|2}$		$\cdots$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$z_t = m_z$	$\alpha_{1\|m_z}$		$\cdots$	$\alpha_{m_c\|m_z}$

where the unknown parameter α is the $(m_c \times m_c)$-dimensional matrix, which exists for each value $k \in \{1, 2, \ldots, m_z\}$ of z_t. Its entries $(\alpha_{i|j})$ are non-negative probabilities of the pointer $c_t = i$ under condition that $z_t = k$.

2.2 Schematic summary of algorithm

The mixture estimation algorithm is obtained via marginalization of the joint pdf for all unknown variables to be estimated. Based on [15, 14, 13, 16], after initialization, the recursive mixture estimation algorithm includes the following steps at each time instant t:

1. Measure new data.
2. Compute the proximity of the current data item to each component. It is done by substituting the measured data item and the point parameter estimates from the previous time instant into corresponding components.
3. Construct the weighting vector containing the probabilities of the activity of components at the actual time instant. It is obtained by using the obtained proximities, the prior pointer pdf and the previous point estimate of α.
4. Declare the active component according the biggest entry of the vector w_t, which is the point estimate of the pointer c_t at time t. Classify the data as belonging to the active component.
5. Update the statistics of all components using the weighting vector, and statistics of the pointer model, using the weighting matrix joint for the current pointer c_t and the previous one c_{t-1}.
6. Recompute the point estimates of all parameters and then use them as the initial ones in the first step of the on-line algorithm.

Here this information is outlined only verbally for better understanding the sense of performed experiments. The detailed description of algorithms based on [15, 14, 13, 16] is available in [17, 18, 19].

3 Comparison experiments

The experiments are performed in the open source programming environment Scilab (www.scilab.org), which presents a powerful tool for engineering computations. All the used codes are completely editable and can be tailored for different specific tasks (including other components or real data application). For all types of components firstly three components were simulated with the dynamic data-dependent pointer model, and one type (normal) was simulated similarly for the static pointer model. Then the mixture-based classification was performed (i) using the same structure of the pointer model and (ii) without data in the condition. The quality of classification was compared using the following criteria:

- Number of incorrect classifications (i.e., the point estimates of the pointer);
- The evolution of individual weights (whether they are close to 1 or 0 for each component);
- The error of the data prediction from the components.

All the experiments are performed for a different level of the model noise, which defines location of components in the data space and a distance among them. The description of experiments is given below for each type of components.

3.1 Categorical components

3.1.1 Data simulation

To perform simulation, parameters for three categorical components (2) and for a pointer model (9) have been set. The total number of samples simulated for three components is 150. After simulation data for the pointer c_t and the discrete variable z_t with three possible values are available. The example of simulation results can be seen in Figure 1.

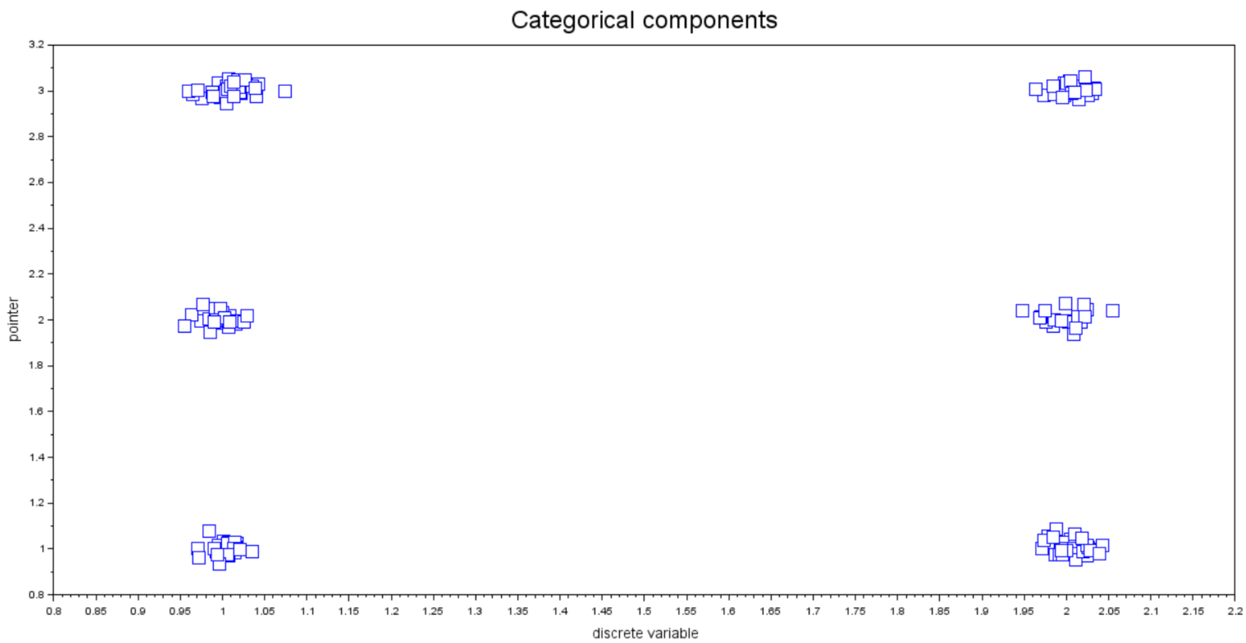


Figure 1: Simulation of categorical components

To make comparison of a data-dependent pointer model (with z_t) and a data-independent pointer model (without z_t) the different degree of uncertainty both for a component model and a pointer model has been used. This uncertainly should be understood as closeness of the model parameters (i.e., probabilities) to the value 1. The closer to the 1, the more deterministic the model is. Such kind of a noise varied in values 0.0001, 0.001, 0.1, 1 and 5 for the component models and 0.001, 0.1, 1 and 5 for the pointer models. Thus, there are 20 different combinations of noise for the component and pointer models. For each combination ten experiments have been performed. The total number of fulfilled experiments is 200.

The results of classification and prediction are presented in the next section.

3.1.2 Results

The results of classification and prediction are summarized for the component models with noise 0.0001, 0.001, 0.1, 1 and 5 and the pointer model with noise 0.001, 0.1, 1 and 5 in Table 1 (classification) and Table 2 (prediction). In the tables the columns are for different variations of noise in a component model. The columns comprise information about classification resp. prediction results for simulations with a data-dependent pointer model (with $z(t)$) and a data-independent pointer model (without $z(t)$). The rows are for the pointer model with noise 0.001, 0.1, 1 and 5. The results are ordered from the smallest number of classification resp. prediction errors for classification resp. prediction with a data-dependent pointer model. Note, that in table with classification results there are two values via '/'. The first value is for wrong classification, the value after '/' is for wrong prediction. It is made to show that the smallest number of incorrectly classified samples does not logically mean the smallest number of incorrectly predicted data.

Table 1: Categorical components: classification/prediction results

Pointer model - noise	Component model - noise									
	0.0001		0.001		0.1		1		5	
	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)
0.001	0/0	36/0	0/0	37/0	7/2	46/83	87/64	96/56	89/0	93/10
	0/0	36/0	1/0	38/149	14/0	43/0	88/11	91/57	89/7	93/108
	1/0	37/0	2/0	34/0	17/0	47/54	91/87	80/0	94/52	90/10
	1/2	24/0	2/0	37/0	23/0	51/47	91/149	103/149	94/68	96/10
	1/4	0/0	2/1	39/51	26/14	47/0	92/1	63/2	97/61	95/1
	1/4	38/149	37/0	37/2	50/1	50/32	92/2	87/6	99/5	98/71
	2/0	29/4	37/0	42/0	60/117	51/3	94/20	99/34	99/57	104/67
	38/149	37/149	38/105	34/0	75/59	53/38	97/13	91/29	101/32	92/2
	75/0	35/10	40/0	38/0	77/7	50/0	109/14	75/58	102/38	104/55
	76/1	38/149	41/1	38/149	78/4	36/10	114/92	94/0	107/58	85/45
0.1	12/0	52/61	11/0	14/6	33/5	53/92	75/19	84/56	86/34	87/24
	22/2	40/6	21/6	29/0	47/6	60/0	76/24	88/149	93/38	92/135
	22/15	36/30	21/4	53/149	49/0	32/0	77/21	95/74	93/79	103/0
	27/11	40/0	35/2	52/2	49/0	30/71	79/0	64/5	94/1	95/1
	36/2	47/2	39/7	39/0	53/0	54/75	81/3	99/67	94/59	94/32
	39/0	45/4	44/2	51/40	60/3	77/16	87/40	99/44	95/34	104/25
	41/2	35/24	44/149	16/0	67/3	41/2	88/36	97/61	98/40	91/81
	41/149	44/1	50/3	50/54	78/24	56/4	88/78	106/1	103/21	98/13
	49/31	41/49	58/0	25/2	81/0	69/0	100/39	84/64	105/2	101/32
	69/2	56/52	67/1	39/10	87/57	48/2	108/48	95/67	105/13	116/22
1	35/1	50/1	32/3	63/40	45/0	79/51	73/2	93/38	87/0	82/20
	44/1	56/0	42/0	50/0	50/1	54/1	76/33	84/33	89/37	97/67
	46/1	47/16	44/1	37/0	56/36	53/7	77/8	76/0	93/67	96/67
	48/0	48/0	45/9	62/43	58/0	51/0	88/33	95/22	94/30	89/57
	49/1	50/1	46/6	47/5	62/7	56/2	93/14	79/15	94/52	102/66
	50/1	42/0	56/5	44/1	63/0	63/0	93/72	93/48	96/69	95/42
	53/2	54/2	56/10	45/0	63/21	71/0	96/77	95/83	100/52	97/54
	54/3	58/2	58/0	46/38	69/47	55/0	98/30	90/0	102/1	101/0
	55/3	51/1	62/0	43/0	76/3	59/0	100/49	90/0	102/23	101/21
	56/3	34/3	63/138	53/1	78/56	65/42	107/66	108/66	114/74	90/66
5	43/22	42/1	42/1	46/1	50/4	59/1	78/10	78/10	84/27	84/80
	44/8	53/27	44/0	50/1	52/0	80/57	81/29	74/4	89/0	89/0
	48/0	50/10	47/0	55/1	55/0	59/0	84/55	91/45	93/57	104/57
	48/0	60/0	50/0	51/2	56/0	54/0	84/69	73/1	94/40	93/16
	49/0	55/17	51/1	50/0	57/0	60/149	84/149	81/149	94/58	103/29
	52/0	35/21	55/0	40/0	59/1	64/0	87/108	82/11	99/37	96/69
	52/0	47/0	56/1	59/16	61/0	63/0	88/72	90/125	100/0	94/70
	52/1	46/0	56/2	56/4	63/0	56/7	90/33	81/43	100/1	106/53
	62/20	45/9	57/0	48/1	68/28	51/0	93/0	92/19	103/65	108/46
	70/71	50/1	62/12	57/1	69/0	69/11	94/64	87/86	110/65	104/0

Table 2: Categorical components: prediction results

Pointer model - noise		Component model - noise									
		0.0001		**0.001**		**0.1**		**1**		**5**	
		with z(t)	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**
	0.001	0	0	0	0	0	0	1	2	0	10
		0	0	0	0	0	47	2	6	5	71
		0	0	0	0	0	54	11	57	7	108
		0	4	0	0	1	32	13	29	32	2
		0	10	0	0	2	83	14	58	38	55
		1	149	0	2	4	10	20	34	52	10
		2	0	0	149	7	0	64	56	57	67
		4	0	1	51	14	0	87	0	58	45
		4	149	1	149	59	38	92	0	61	1
		149	149	105	0	117	3	149	149	68	10
	0.1	0	4	0	2	0	0	0	5	1	1
		0	61	0	6	0	0	3	67	2	32
		2	2	1	10	0	71	19	56	13	22
		2	6	2	2	0	75	21	74	21	13
		2	24	2	40	3	2	24	149	34	24
		2	52	3	54	3	16	36	61	34	25
		11	0	4	149	5	92	39	64	38	135
		15	30	6	0	6	0	40	44	40	81
		31	49	7	0	24	4	48	67	59	32
		149	1	149	0	57	2	78	1	79	0
	1	0	0	0	0	0	0	2	38	0	20
		1	0	0	0	0	0	8	0	1	0
		1	0	0	38	0	51	14	15	23	21
		1	1	1	0	1	1	30	0	30	57
		1	1	3	40	3	0	33	22	37	67
		1	16	5	1	7	2	33	33	52	54
		2	2	6	5	21	0	49	0	52	66
		3	1	9	43	36	7	66	66	67	67
		3	2	10	0	47	0	72	48	69	42
		3	3	138	1	56	42	77	83	74	66
	5	0	0	0	0	0	0	0	19	0	0
		0	0	0	1	0	0	10	10	0	70
		0	10	0	1	0	0	29	4	1	53
		0	17	0	1	0	7	33	43	27	80
		0	21	0	2	0	11	55	45	37	69
		1	0	1	0	0	57	64	86	40	16
		8	27	1	1	0	149	69	1	57	57
		20	9	1	16	1	0	72	125	58	29
		22	1	2	4	4	1	108	11	65	0
		71	1	12	1	28	0	149	149	65	46

Further on, the figures illustrating outputs for some component/pointer model noise variations are presented. Simulation results for the component model with noise 0.0001 and the pointer model with noise 0.001 are shown in Figure 2.

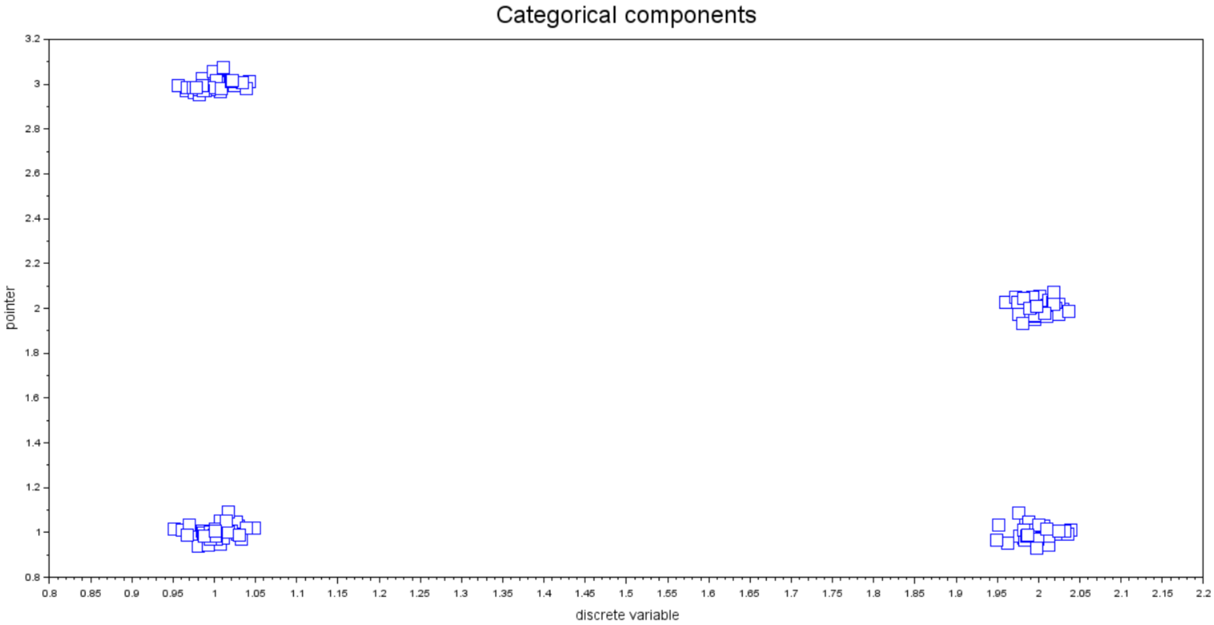
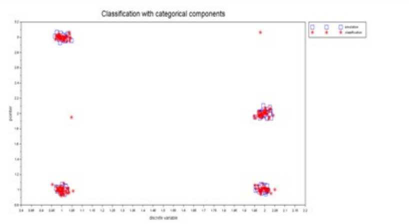
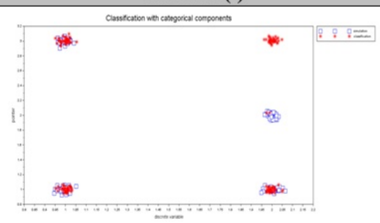
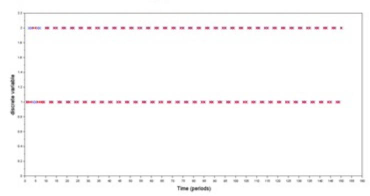
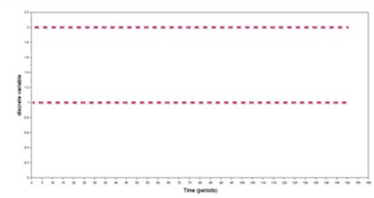
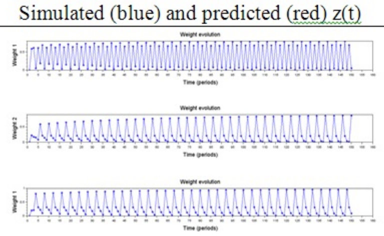
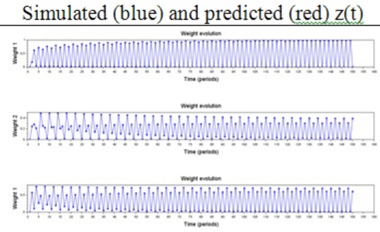


Figure 2: Simulation of categorical components: component model with noise 0.0001/pointer model with noise 0.001

In the following table the results of classification with $z(t)$ and without $z(t)$ are listed.

Table 3: Comparison of simulation results with a data-dependent and a data-independent pointer model

with z(t)	without z(t)
 Classification/prediction error: 1/6	 Classification/prediction error: 36/0
 Simulated (blue) and predicted (red) z(t)	 Simulated (blue) and predicted (red) z(t)
	

The same is made for the component models with noise value 0.0001 and the pointer model with noise value 5, as well as for the component models with noise value 5 and the pointer models with noise values 0.001 and 5. The results can be seen in the following figures and tables.

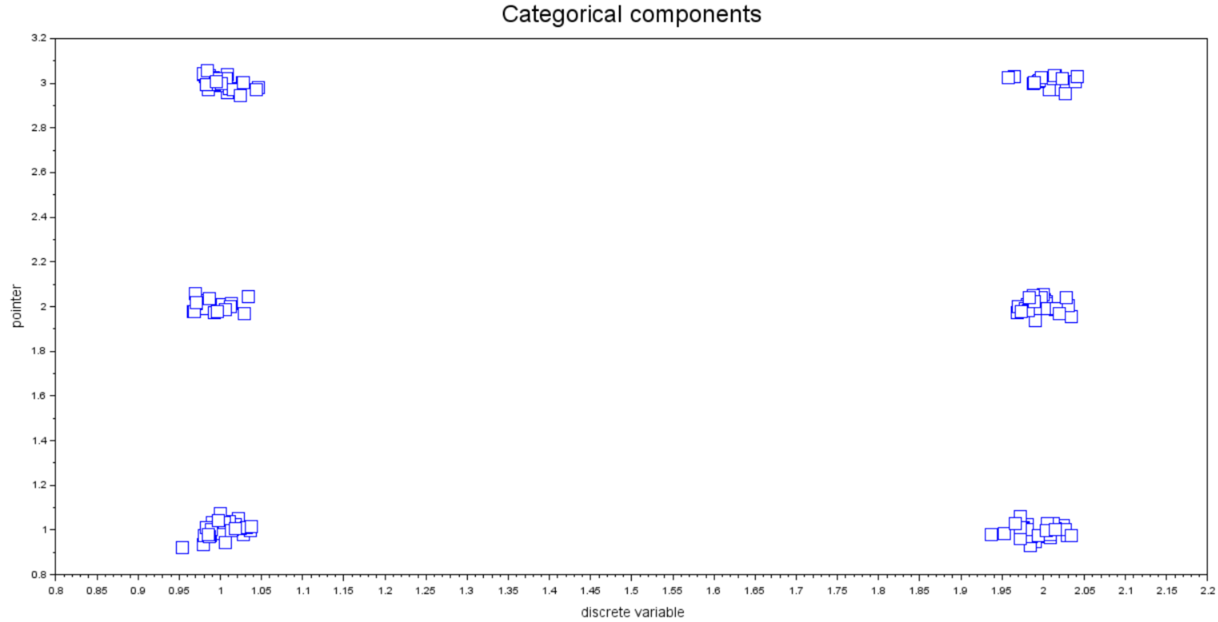
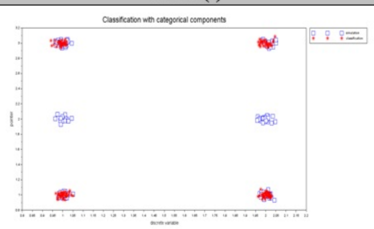
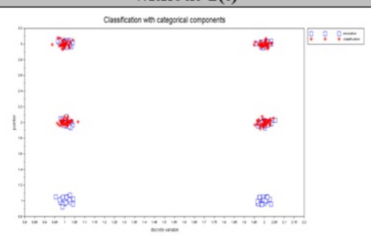
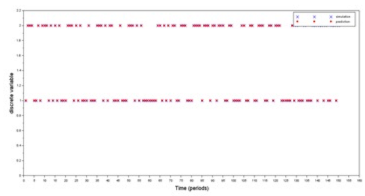
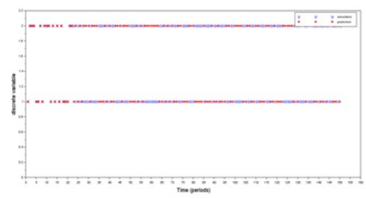
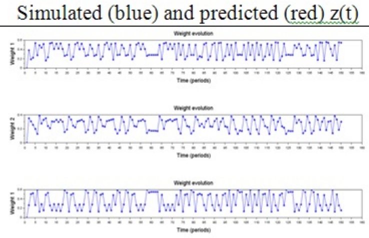
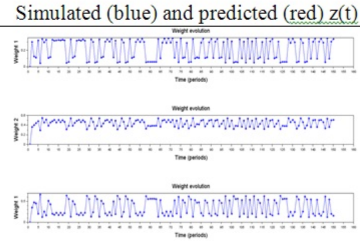


Figure 3: Simulation of categorical components: component model with noise 5/pointer model with noise 0.001

Table 4: Comparison of simulation results with a data-dependent and a data-independent pointer model

with z(t)	without z(t)
 Classification/prediction error: 88/0	 Classification/prediction error: 99/128
 Simulated (blue) and predicted (red) z(t)	 Simulated (blue) and predicted (red) z(t)
	

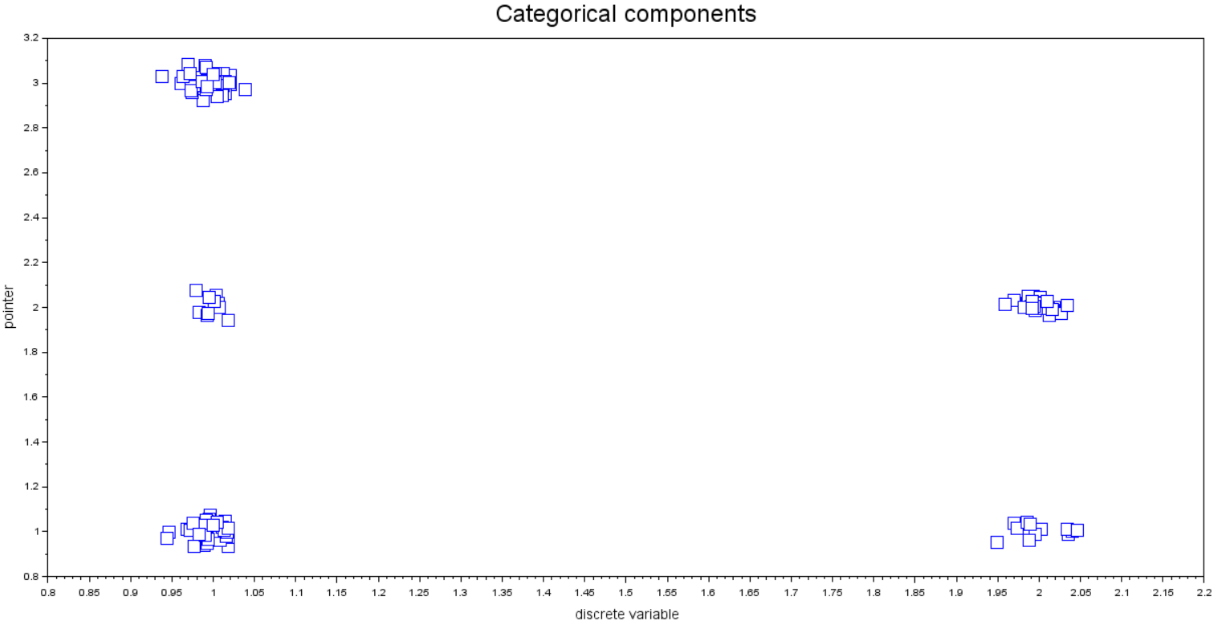
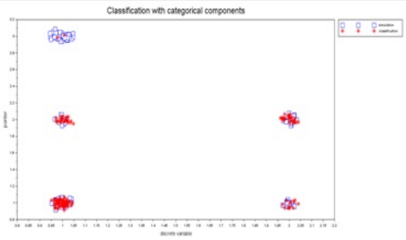
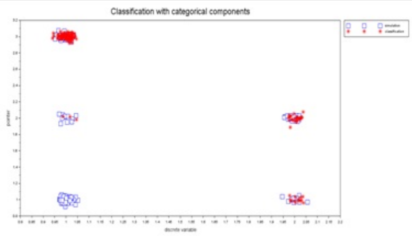
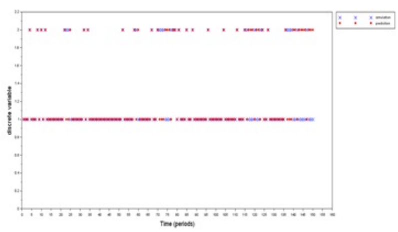
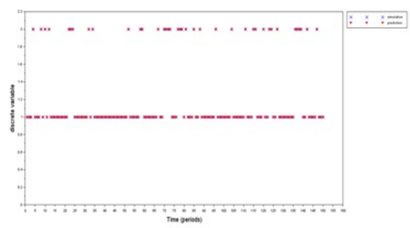
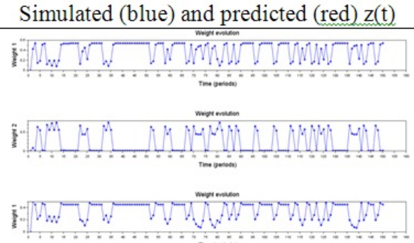
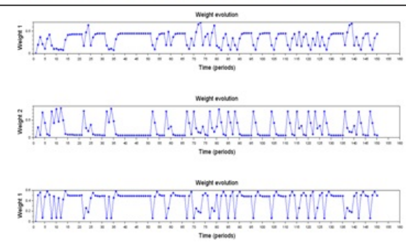


Figure 4: Simulation of categorical components: component model with noise 0.0001/pointer model with noise 5

Table 5: Comparison of simulation results with a data-dependent and a data-independent pointer model

with z(t)	without z(t)
 Classification/prediction error: 62/34	 Classification/prediction error: 56/149
 Simulated (blue) and predicted (red) z(t)	 Simulated (blue) and predicted (red) z(t)
	

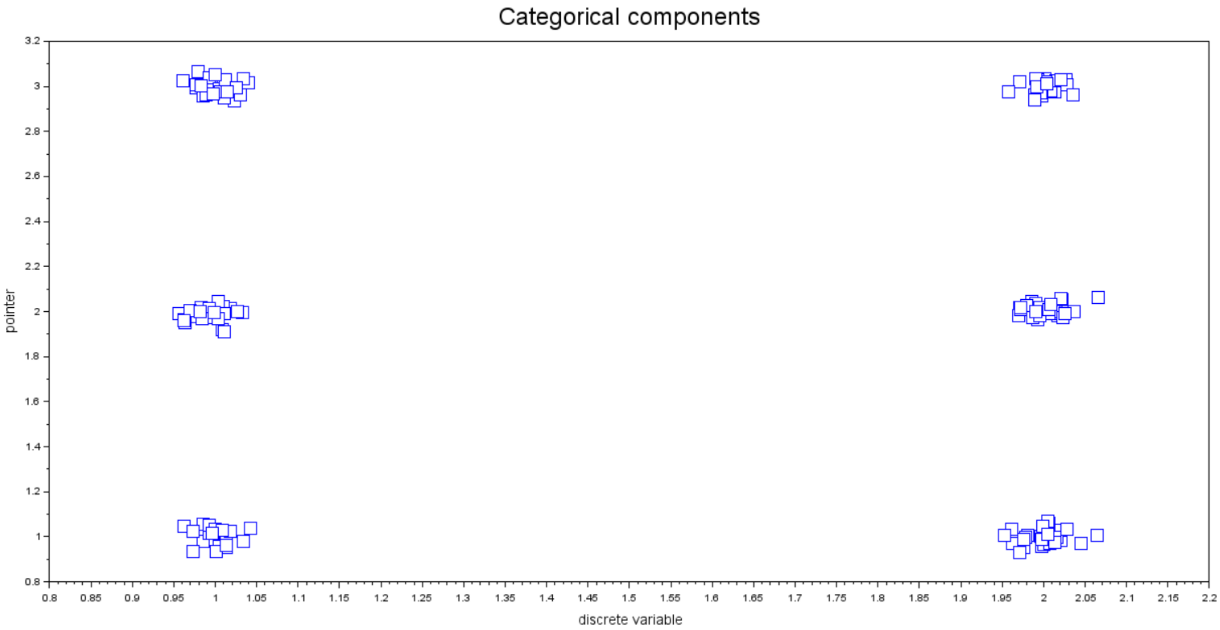
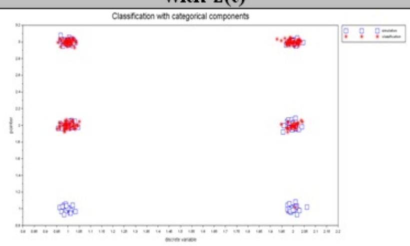
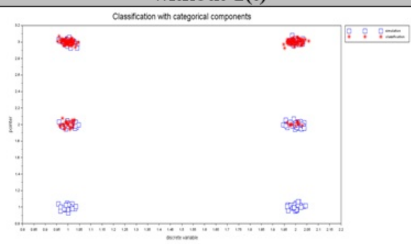
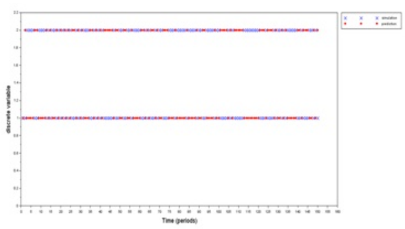
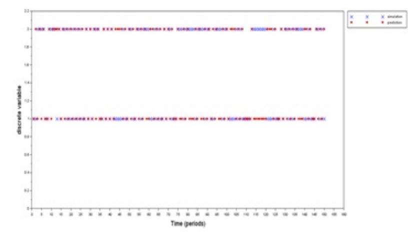
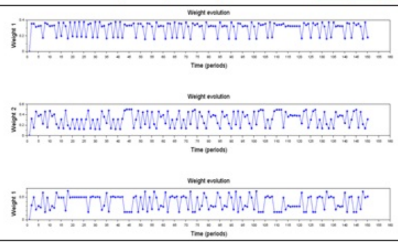
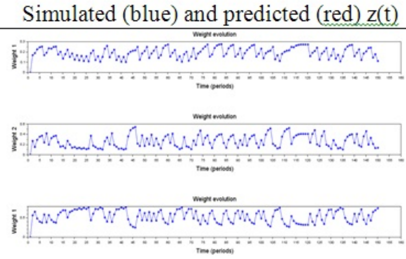


Figure 5: Simulation of categorical components: component model with noise 5/pointer model with noise 5

Table 6: Comparison of simulation results with a data-dependent and a data-independent pointer model

with z(t)	without z(t)
 Classification/prediction error: 91/149	 Classification/prediction error: 100/91
 Simulated (blue) and predicted (red) z(t)	 Simulated (blue) and predicted (red) z(t)
	

3.1.3 Discussion

According to the results listed in tables the number of incorrectly classified samples is the smallest one for noise values in the component models 0.0001 and 0.001 while the noise of pointer model is 0.001. The higher a noise value is the greater number of incorrectly classified samples are presented in the model. The largest number of incorrectly classified data is when the noise value is 5 both for the component and pointer models. It is valid both for data-dependent and data-independent pointer models.

The number of wrong classifications using a data-dependent model for a component model with noise value 0.0001 and 0.001 and for a pointer model with noise value 0.001 varies from 0 to 76 samples (i.e. 0 % to 51.01 %) and 0 to 41 samples (0 % to 27.52 %) respectively. The average number of wrong classifications is 19.5 (13.09 %) and 20 (13.42 %) resp., while standard deviation for the first case is very large - 31.71 samples (21.28 %). For the second case the standard deviation is 19.65 samples (13.19 %).

As for the data-independent model the number of wrong classifications is a little bit larger for the same values of noise: from 0 to 38 samples (0 % to 25.50 %) and from 34 to 42 samples (22.82 % to 28.19 %) resp. The average values are 31 samples (20.81 %) and 37.40 samples (25.10 %) resp. And the standard deviation is 11.79 samples (7.91 %) and 2.32 samples (1.56 %).

It should be noted that though in two discussed cases the average number of incorrectly classified samples is smaller for a data-dependent pointer model, however, the standard deviation is significantly higher and not in all experiments performed for these noise values the data-dependent model shows the best results. There are cases, for example, when the number of wrong classifications is twice greater for a data-dependent pointer model than for a data-independent pointer model: 75 via 35 samples or 76 via 38 samples (see table below).

Besides, if taking into account all experiments we see that there is no evident prove that a data-dependent model performs better, as in many cases a data-independent model gives better or similar results in classification. In the following table the results are presented in more details. Green color in the table means success of a data-dependent pointer model over a data-independent pointer model. Red color rows are for the greater number of wrong classifications using a data-dependent pointer model. The equal number of wrong classifications for both models are colored in blue. In the table the average and standard deviations are presented.

Table 7: Categorical components: prediction results

		Component model - noise									
		0.0001		0.001		0.1		1		5	
		with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)
Pointer model - noise	0.001	19.50±31.71	31±11.79	20±19.65	37.40±2.32	42.70±28.37	47.40±4.97	95.50±8.96	87.90±12.12	97.10±5.72	95±5.91
	0.1	35.80±16.21	43.60±6.69	39±17.52	36.80±15.02	60.40±17.42	52±14.98	85.90±10.88	91.10±11.82	96.60±6.13	98.10±8.33
	1	49±6.31	49±6.99	50.40±10.07	49±8.27	62±10.48	55.60±20.08	90.10±11.38	90.30±9.09	97.10±7.80	95±6.32
	5	52±8.23	48.30±7.02	52±6.32	51.20±5.75	59±6.32	61.50±8.29	86.30±5.10	82.90±6.87	96.60±7.40	98.10±8.05

As far as prediction is concerned the better results are shown by a data-dependent pointer model in the majority of cases with different noise values. It should be noted that standard deviation, however, is very high.

Table 8 lists the outcomes of experiments. It comprises average values of incorrect predictions and their standard deviations. The greater number of incorrectly predicted samples using a data-dependent pointer model in comparison with a data-independent pointer model is in red. Green rows mean the less number of wrong predictions made by a data-dependent pointer model.

Table 8: Prediction: average and standard deviation

Pointer model - noise	Component model - noise: 0.0001		0.001		0.1		1		5	
	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)
0.001	16±46.76	46.10±71.08	10.70±33.14	35.10±62.08	20.40±38.41	26.70±28.80	45.30±50.22	39.10±45.58	37.80±25.64	37.90±36.80
0.1	21.40±45.88	22.90±23.84	17.40±46.30	26.30±47.08	9.80±18.09	26.20±37.34	30.80±22.87	58.80±40.82	32.10±24.47	36.50±41.32
1	1.60±1.08	2.60±4.81	17.20±42.61	12.80±19.10	17.10±21.65	10.30±19.32	38.40±26.81	30.50±28.85	40.50±26.94	46±24.27
5	12.20±22.37	8.60±10.01	1.70±3.68	2.70±4.81	3.30±8.77	22.50±47.79	58.90±45.05	49.30±53.11	35±26.90	42±29.24

3.2 Uniform components

3.2.1 Data simulation

Before making classification the parameters for the pointer model and two parameters of the component model - a and b, the minimum and maximum values, are defined. For three components the total number of simulated data is set to 500, including data for component model $y(t)$, data for pointer model $c(t)$, discrete values for a pointer model $z(t)$ and weights α. The data are saved using SciLab function "save" and then used for estimation for both a data-dependent and data-independent pointer model.

During simulation we have changed the values of parameters a and b by multiplying corresponding matrices by 0.8, 1, 1.8, 3 and 10; thus, making distance between components greater or smaller. The higher a multiplying value of matrix b is the nearer the components of the component model are and the harder to perform satisfactory estimation and classification is. The multiplication of matrix a with higher values results in more compact data location in the component model. The multiplication was made turn by turn, so there are 25 combinations for different values of a and b. For each combination ten experiments were performed. The total number of experiments for uniform components is 250.

Some examples of simulated data can be seen in the following figures.

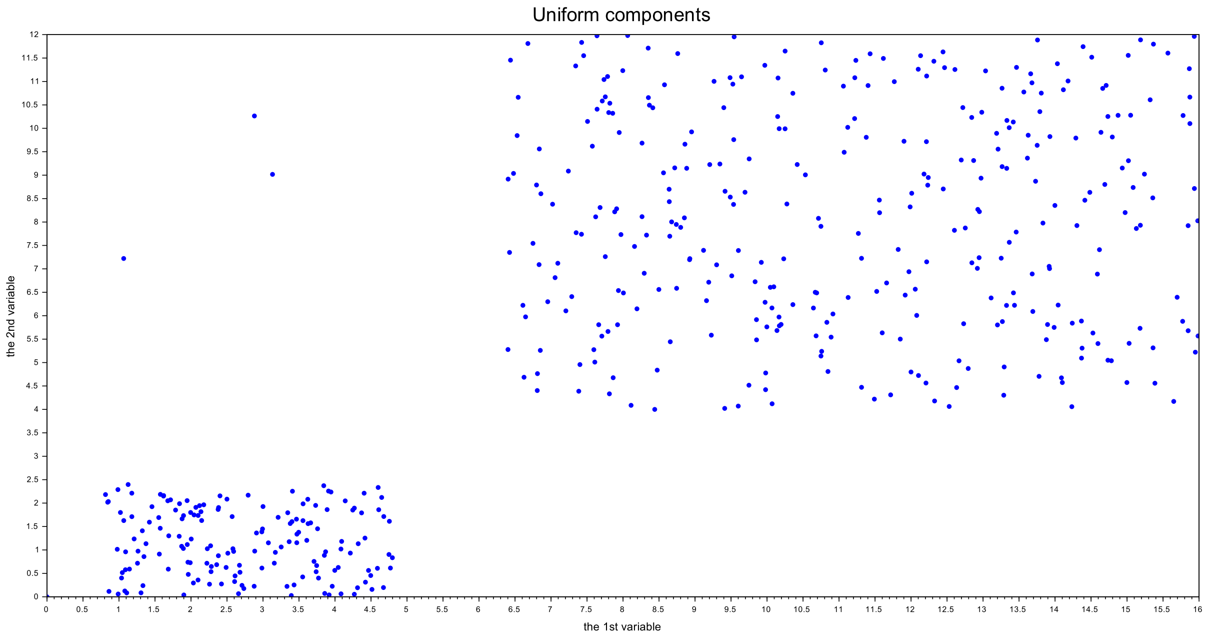


Figure 6: Simulation of uniform components: a and b matrices are multiplied by 0.8

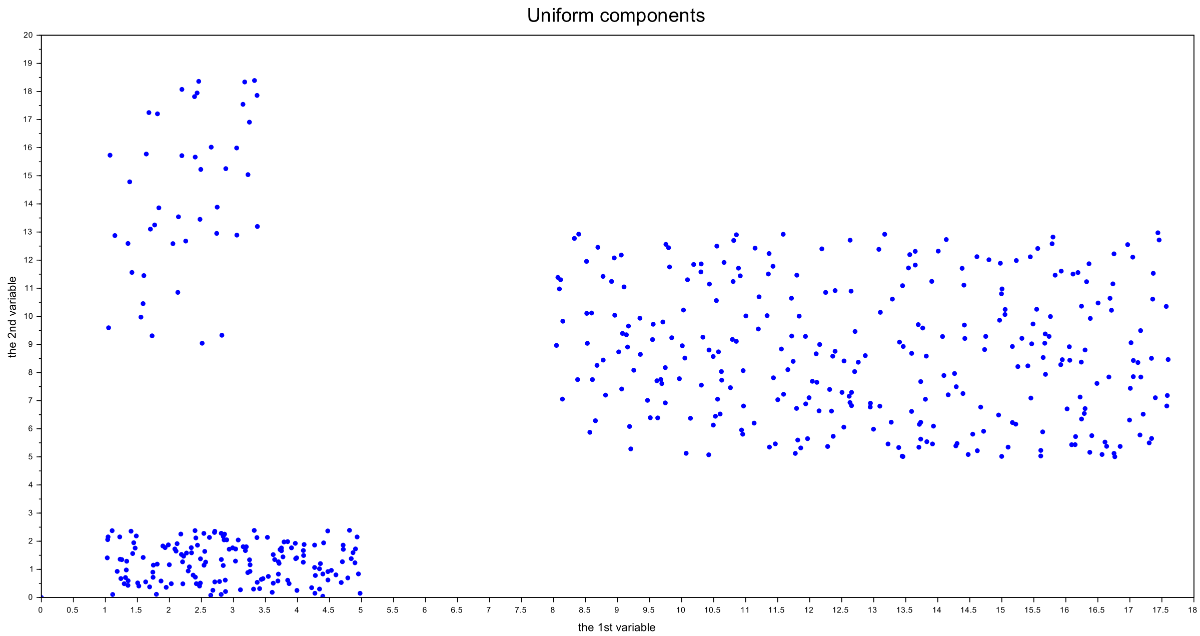


Figure 7: Simulation of uniform components: a is multiplied by 1, b is multiplied by 0.8

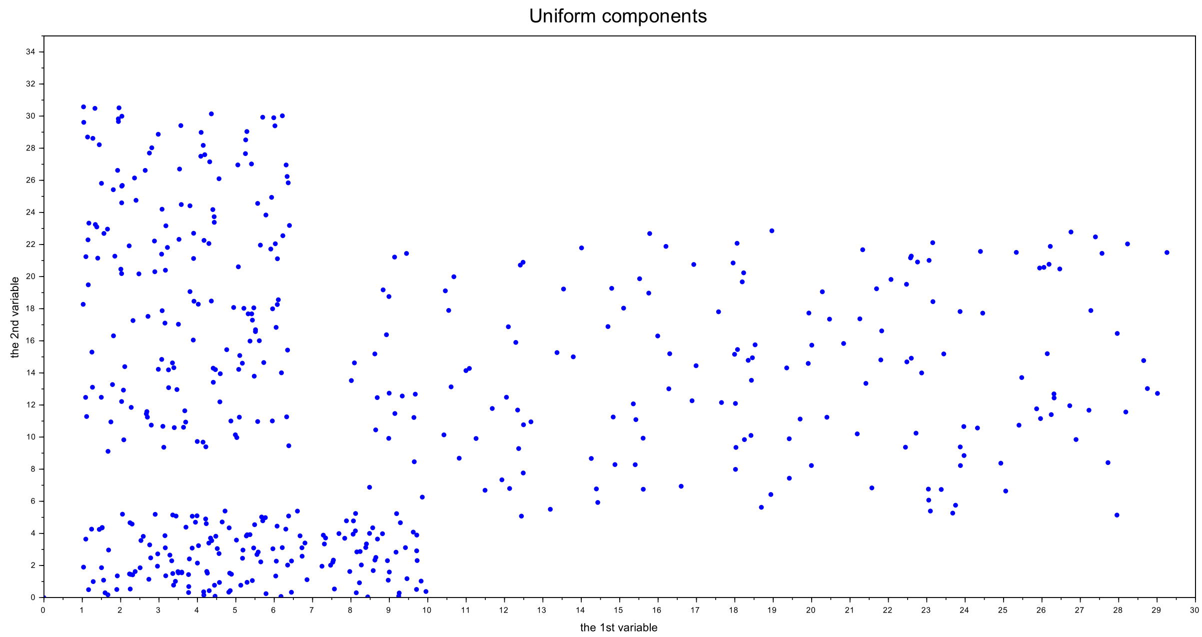


Figure 8: Simulation of uniform components: a is multiplied by 1, b is multiplied by 1.8

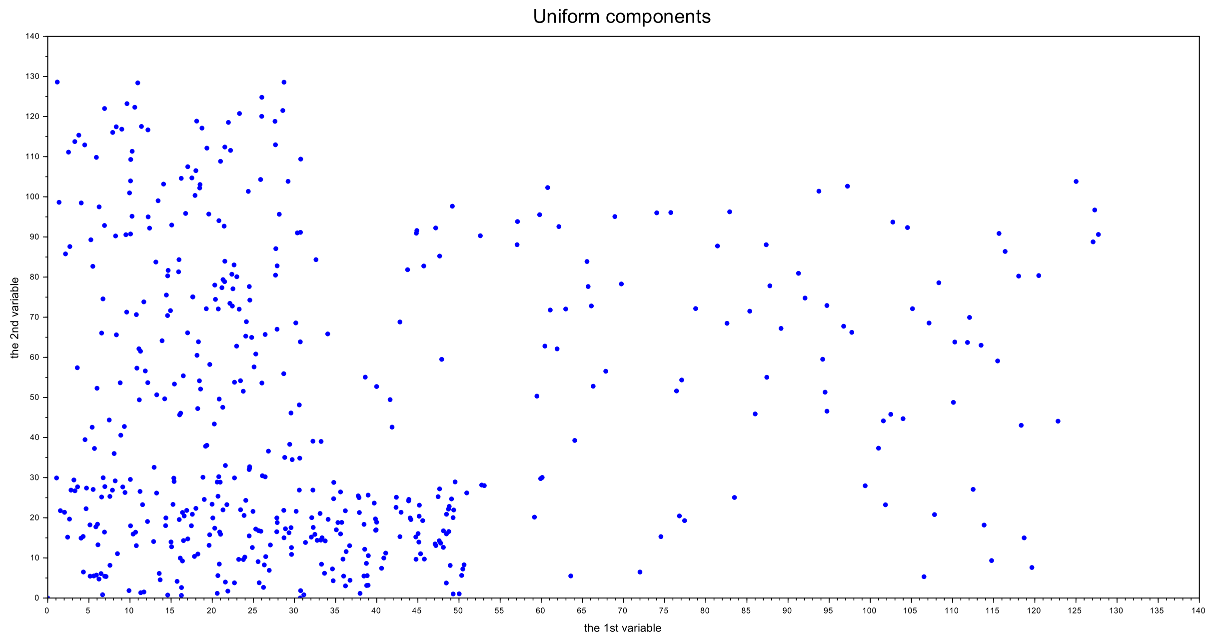


Figure 9: Simulation of uniform components: a is multiplied by 1, b is multiplied by 10

The initial weight vector [0.4, 0.3, 0.3] as well as standard deviation 0.5 for proximity are defined for both models. The initial statistics and estimation for components and the pointer are performed prior to estimation. The results of algorithm are presented in the next section.

3.2.2 Results

The results of simulation and estimation both for the data-dependent and data-independent pointer models are presented in the following table. The table shows the number of incorrectly classified samples with a data-dependent pointer (with $z(t)$) via the number of misclassifications with a data-independent pointer (without $z(t)$). The columns are for values of parameter b and the rows are for the values of parameter a.

Table 9: Classification results for uniform components

	b*0.8		**b*1**		**b*1.8**		**b*3**		**b*10**	
	with z(t)	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**
a*0.8	7	8	16	17	49	50	48	48	82	82
	13	14	29	30	51	52	59	59	101	101
	25	34	29	31	57	57	59	59	101	101
	27	27	43	43	62	62	63	64	115	115
	37	41	45	45	63	64	64	64	118	118
	49	49	46	47	64	64	68	68	129	129
	51	52	49	50	70	70	73	73	135	135
	60	62	55	56	79	79	83	83	273	273
	64	66	60	60	85	85	85	84	300	300
	70	71	65	68	250	250	91	92	341	341
a*1	0	0	1	3	42	42	37	37	90	90
	0	0	2	2	44	44	37	37	96	96
	0	0	8	8	44	44	39	39	136	136
	0	0	8	8	50	50	39	40	145	145
	0	0	9	15	50	50	60	60	172	172
	1	1	12	12	52	52	65	65	226	226
	1	1	12	15	53	53	74	75	239	239
	1	1	14	14	53	53	77	77	265	265
	3	3	17	18	54	54	78	78	267	267
	4	4	42	44	66	66	149	149	353	353
a*1.8	0	0	0	0	1	1	14	14	71	71
	0	0	0	0	1	1	15	15	76	76
	0	0	0	0	1	1	18	18	78	78
	0	0	0	0	1	1	19	19	93	93
	0	0	0	0	3	3	20	20	97	97
	0	0	0	0	10	10	21	21	103	103
	0	0	0	0	10	10	28	28	133	133
	0	0	0	0	11	11	29	29	170	170
	0	0	0	0	13	13	34	34	189	189
	0	0	0	0	13	13	76	77	228	228
a*3	0	0	0	0	0	0	1	1	59	59
	0	0	0	0	0	0	2	2	62	62
	0	0	0	0	0	0	2	2	64	64
	0	0	0	0	0	0	7	7	67	67
	0	0	0	0	0	0	9	9	82	82
	0	0	0	0	0	0	21	21	90	89
	0	0	0	0	0	0	26	26	175	175
	0	0	0	0	3	3	33	33	285	285
	0	0	0	0	4	4	36	36	292	292
	0	0	0	0	4	4	66	66	358	358
a*10	166	166	166	166	179	179	185	185	50	50
	166	166	166	166	179	179	189	189	81	81
	166	166	166	166	182	182	189	189	218	215
	166	166	166	166	182	182	190	190	292	292
	166	166	166	166	182	182	191	191	300	300
	167	167	166	166	183	183	192	192	306	306
	167	167	167	167	183	183	192	192	314	314
	167	167	167	167	183	183	193	193	314	314
	167	167	167	167	183	183	194	194	327	327
	167	167	167	167	183	183	197	197	343	343

Some examples of simulation and estimation are shown in the following figures and tables.
The first example is for multiplication of parameter a with value 1 and parameter b with value 0.8.

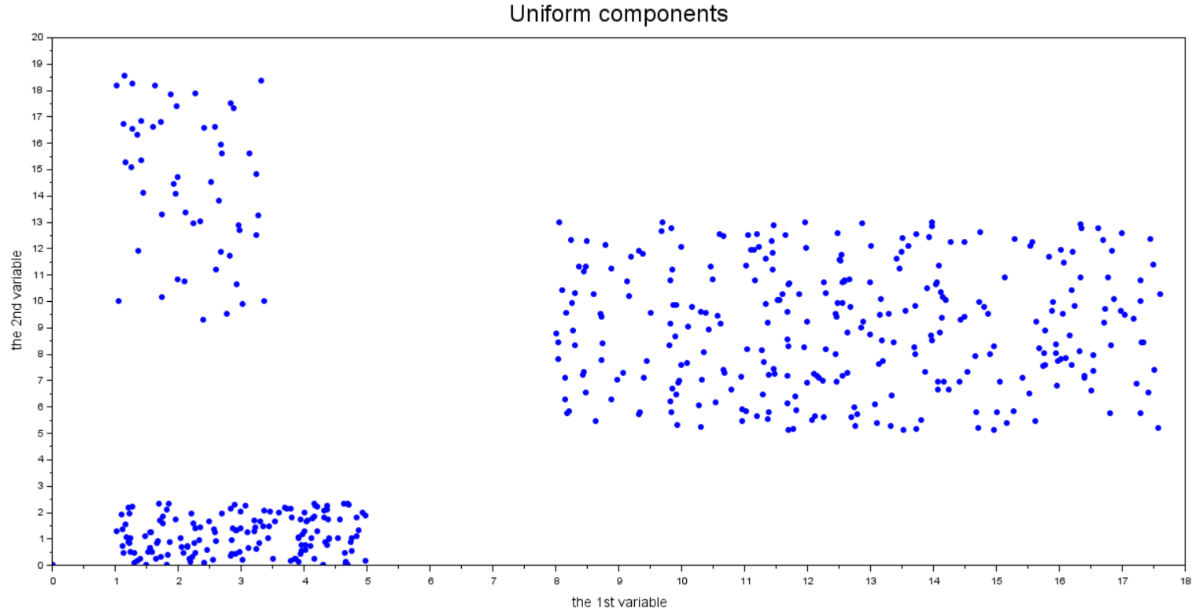
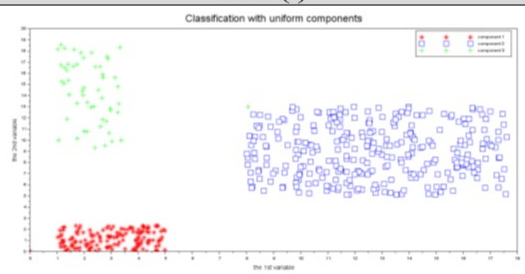
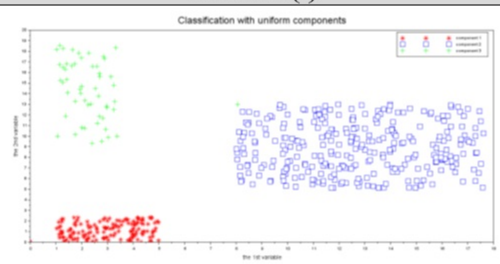
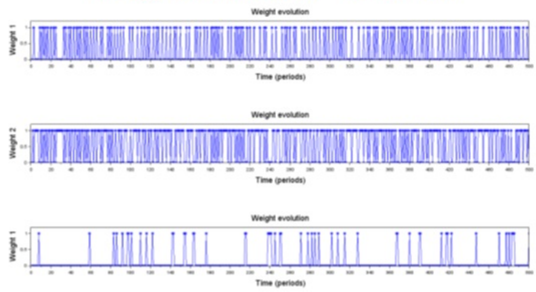
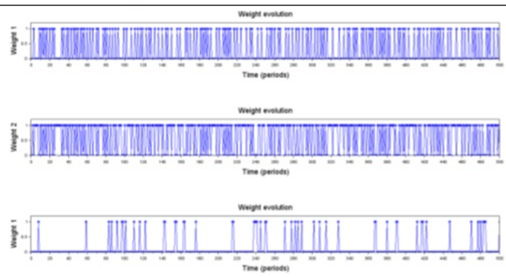


Figure 10: Simulation of uniform components: $a * 1$, $b * 0.8$

From the picture it is clearly seen that three components are far from each other and there should not be difficulties in their classification. The results of estimation and classification both for the data-dependent and data-independent pointer models are presented in the following table.

Table 10: Simulation results: $a * 1$, $b * 0.8$

with z(t)	without z(t)
 Wrong classifications = 1 from 500	 Wrong classifications = 1 from 500
	

The next example is for multiplication values 1 for a and 10 for b.

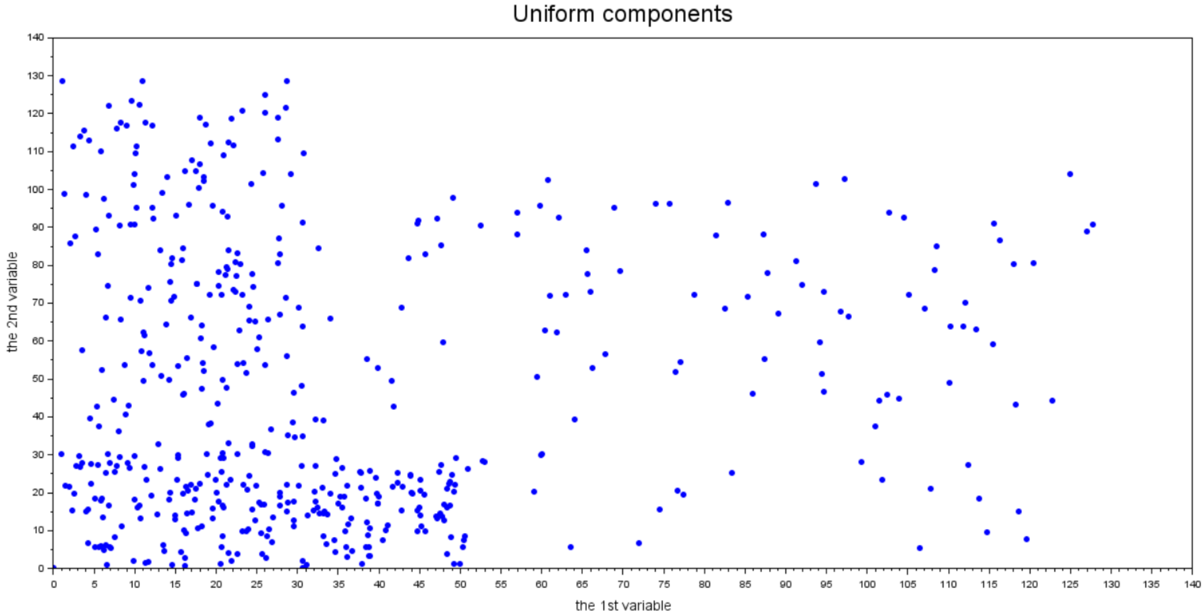
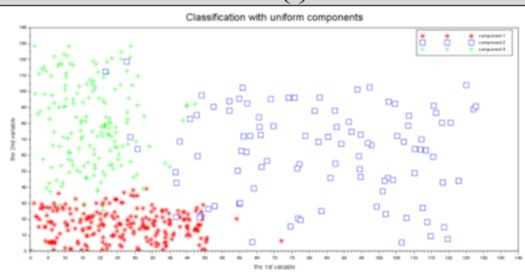
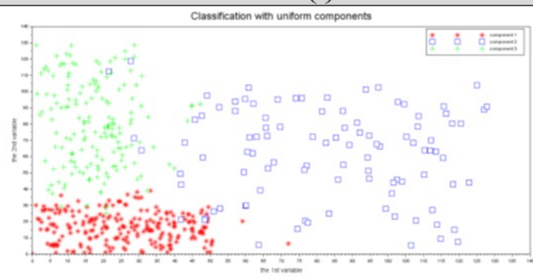
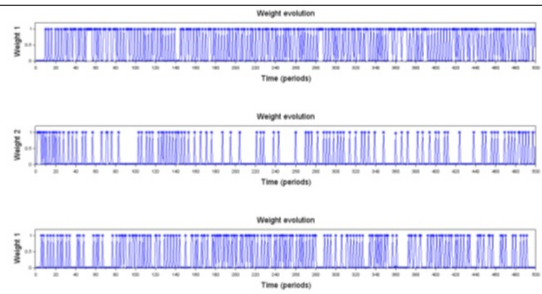
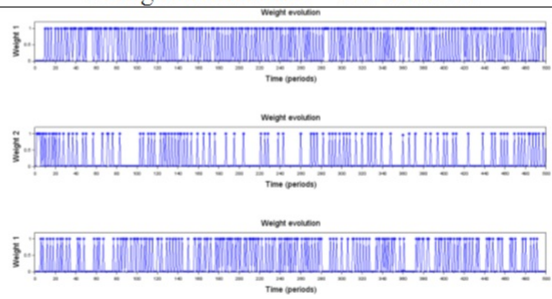


Figure 11: Simulation of uniform components: $a * 1$, $b * 10$

The components are very close to each other. The results of classification are listed in table 11.

Table 11: Classification results: $a * 1$, $b * 10$

with z(t)	**without z(t)**
 Wrong classifications = 100 from 500	 Wrong classifications = 100 from 500
	

Simulation with parameters a and b multiplied by 10 and 1 resp. is shown below.

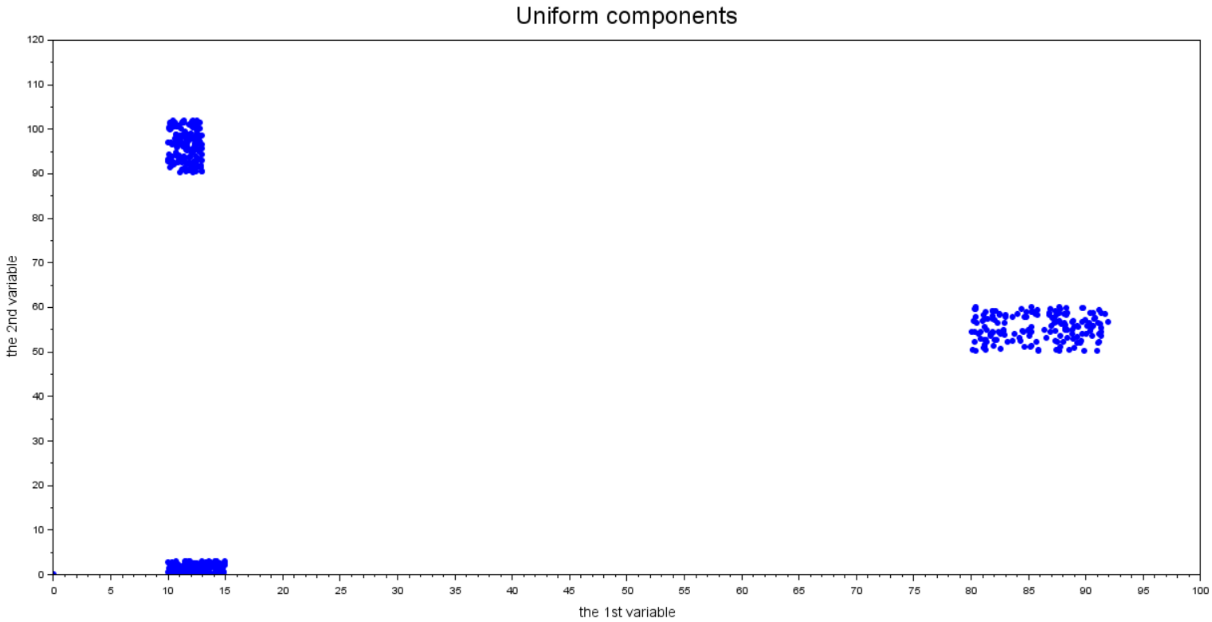
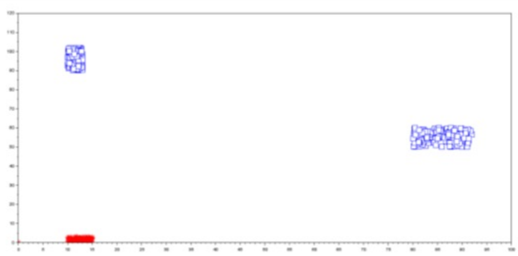
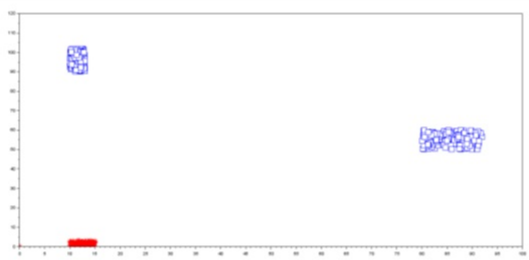


Figure 12: Simulation of uniform components: $a * 10$, $b * 1$

Table 12: Classification results: $a * 10$, $b * 1$

with z(t)	without z(t)
 Wrong classifications = 166 from 500	 Wrong classifications = 166 from 500

Because the third component was not estimated, the programme gave error and did not generate the graph of weight evolution.

The last example presented here is for a and b multiplied by 10.

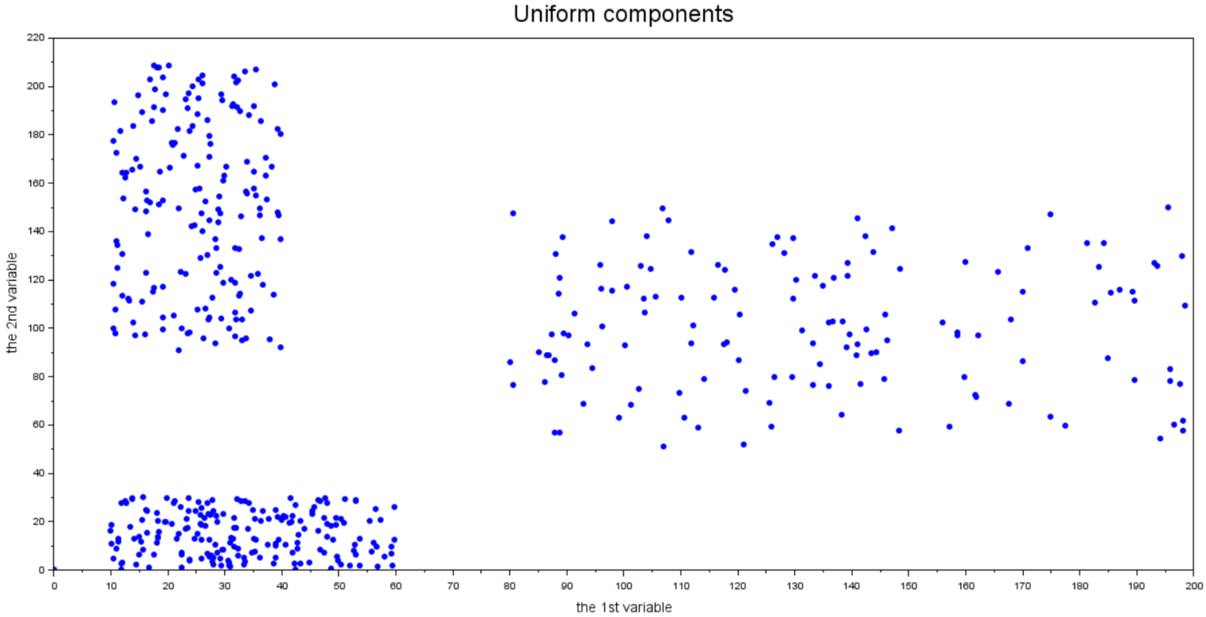
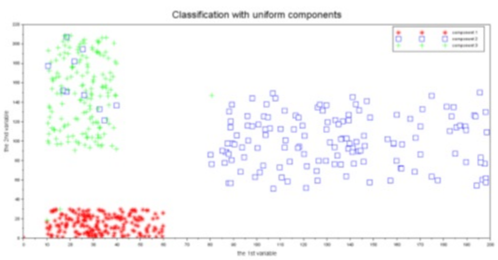
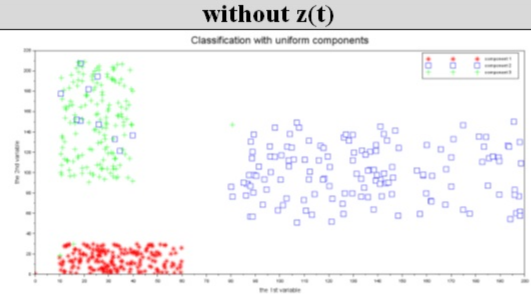
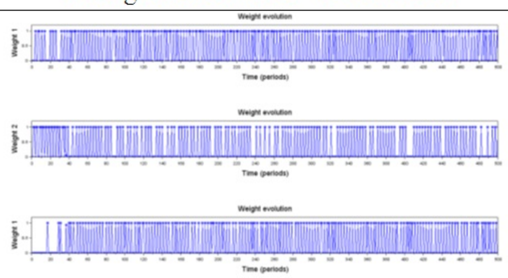
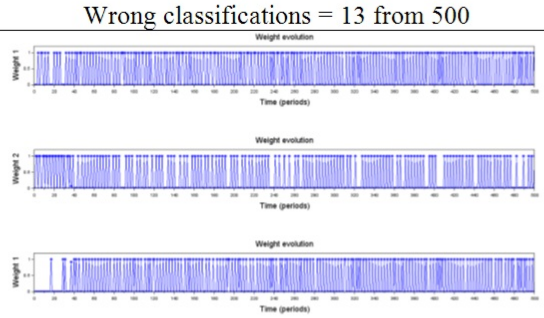


Figure 13: Simulation of uniform components: $a * 10$, $b * 10$

Table 13: Classification results: $a * 10$, $b * 10$

with z(t)	**without z(t)**
 Wrong classifications = 13 from 500	 Wrong classifications = 13 from 500
	

The results of estimation and classification are discussed in Section 3.2.3.

3.2.3 Discussion

From table 9 in the previous section it is obvious that the greater a value of parameter b is, i.e. the farther the components from each other are, the more precise the classification is. The higher values of parameter a also contribute to good performance of the algorithm, as the components become more compact and also are far away from each other. However, it is not valid for $a * 10$, where the number of wrong classifications is quite high.

If we compare the results while using a data-dependent and data-independent pointer model for $a*0.8$ and $b*0.8$, we see that the number of misclassifications is from 7 to 70 samples (1.4 % to 14 %) from total number of 500 samples for a pointer model with $z(t)$ and from 8 to 71 samples from total number of 500 samples (1.6 % to 14.2 %) for a pointer model without $z(t)$. For $a*1.8$ and $b*0.8$ the number of incorrectly classified samples is 0. Thus, there is no much difference in using a pointer model with $z(t)$ and without $z(t)$. The same situation is for all other cases of parameter multiplications.

To make the results clearer table 14 shows the average value and standard deviation for classification results with different values of parameters a and b both for the data-dependent and data-independent pointer models. Green cells in the table mean that the number of incorrectly classified samples is less for a data-dependent model. The equal number of wrong classifications for models with $z(t)$ and without $z(t)$ is highlighted in blue colour. The greater number of misclassifications with a data-dependent model in comparison with a data-independent model is in red colour. From this we see that a data-dependent pointer model performs slightly better, but in the majority of cases the performance of both models can be evaluated as more or less similar. The same can be seen also from the figures in the previous section.

Table 14: Classification results: average and standard deviation

	b*0.8		**b*1**		**b*1.8**		**b*3**		**b*10**	
	with z(t)	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**
a*0.8	40.30±21.85	42.40±21.59	43.70±15.21	44.70±15.32	83±59.76	83.30±59.60	69.30±13.56	69.40±13.57	169.50±95.83	169.50±95.83
a*1	1±1.41	1±1.41	12.50±11.49	13.90±11.83	50.80±6.86	50.80±6.86	65.50±33.98	65.7±33.93	198.90±85.07	198.90±85.07
a*1.8	0	0	0	0	6.40±5.40	6.40±5.40	27.40±18.25	27.50±18.54	123.80±54.36	123.80±54.36
a*3	0	0	0	0	1.10±1.79	1.10±1.79	20.30±20.75	20.30±20.75	153.40±115.81	153.30±115.87
a*10	166.50±0.53	166.50±0.53	166.40±0.52	166.40±0.52	181.90±1.60	181.90±1.60	191.20±3.26	191.20±3.26	254.50±105.17	254.20±105.29

3.3 Exponential components

3.3.1 Data simulation

The first step before the very estimation and classification was to define the parameters of the pointer model and parameters of the component model a and b. For three components 500 samples were simulated including data for a component model $y(t)$, data for a pointer model $c(t)$, discrete variables for a pointer model $z(t)$ and weights α.

The distance between components can be changed by varying parameters of the component model - a and b in this case. The greater the values of the parameters are the farther the components from each other are located. To vary the distance the matrices a and b were multiplied by values 0.01, 0.1, 0.2, 0.5, 1 and 1.5 turn by turn. Thus, there are 36 combinations. For each combination ten experiments are performed. The total number of experiments for exponential components are 360.

Several examples of simulated exponential components are shown in the following figures. The results are presented and discussed in Sections 3.3.2 and 3.3.3 respectively.

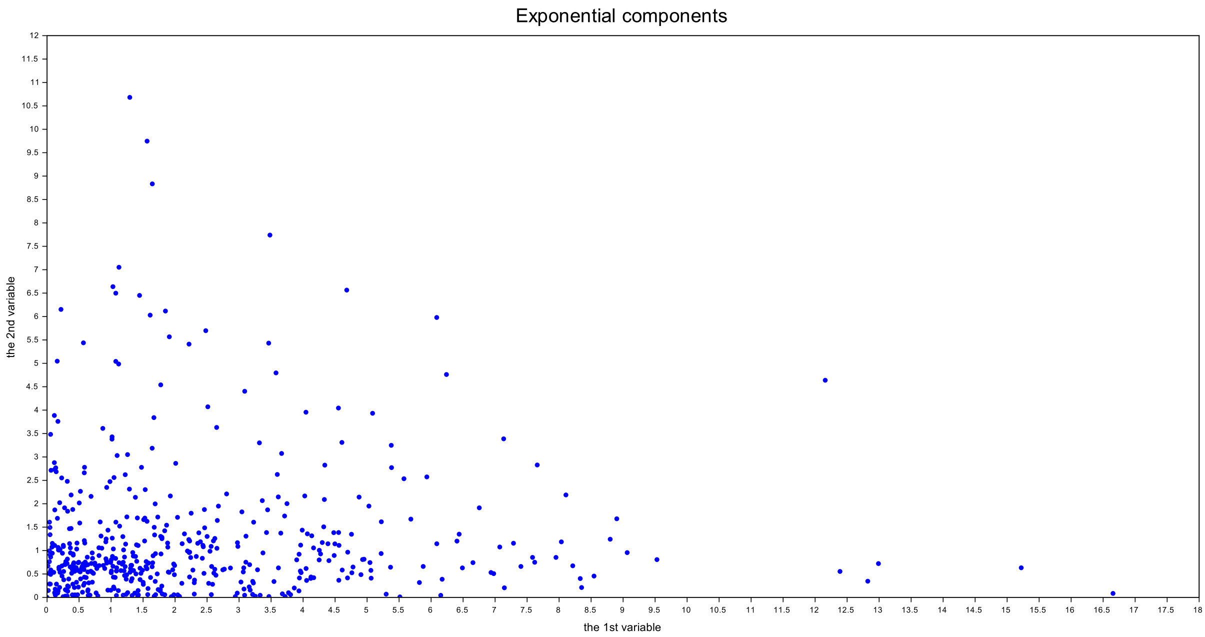


Figure 14: Simulation of exponential components: $a * 0.5$, $b * 0.01$

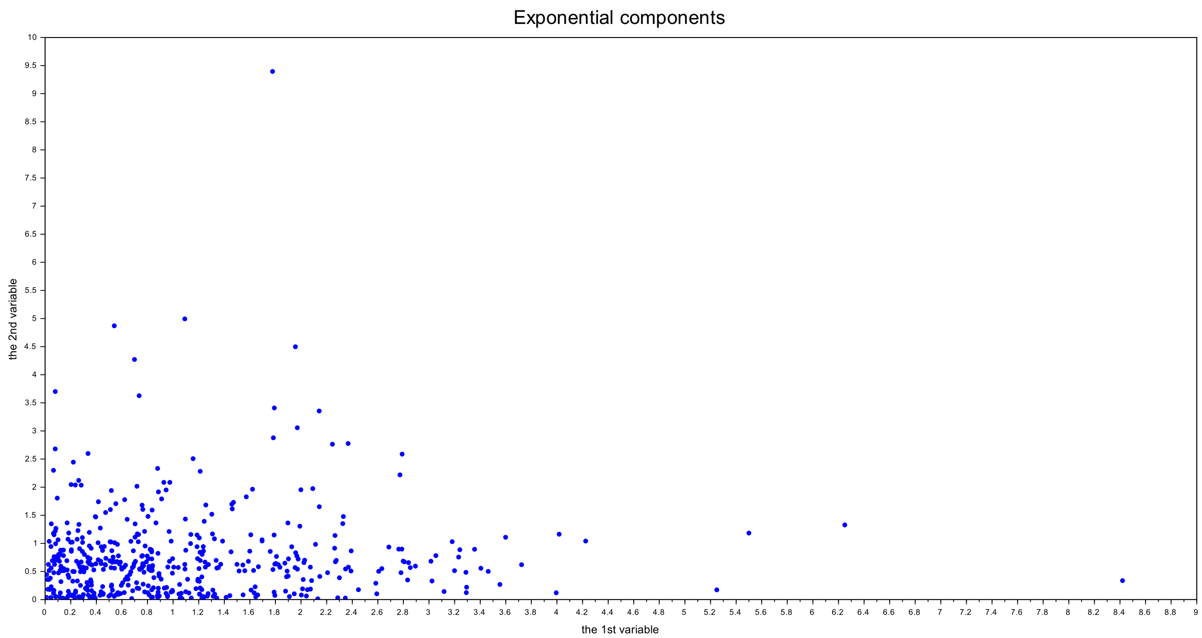


Figure 15: Simulation of exponential components: $a * 1$, $b * 0.01$

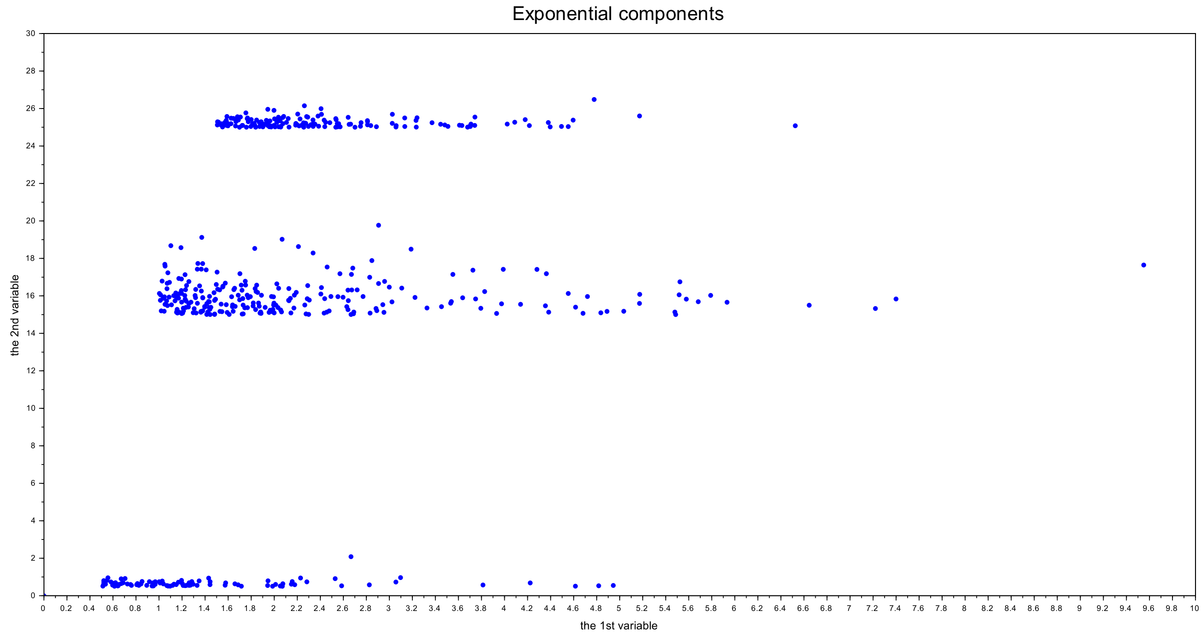


Figure 16: Simulation of exponential components: $a * 0.5$, $b * 0.5$

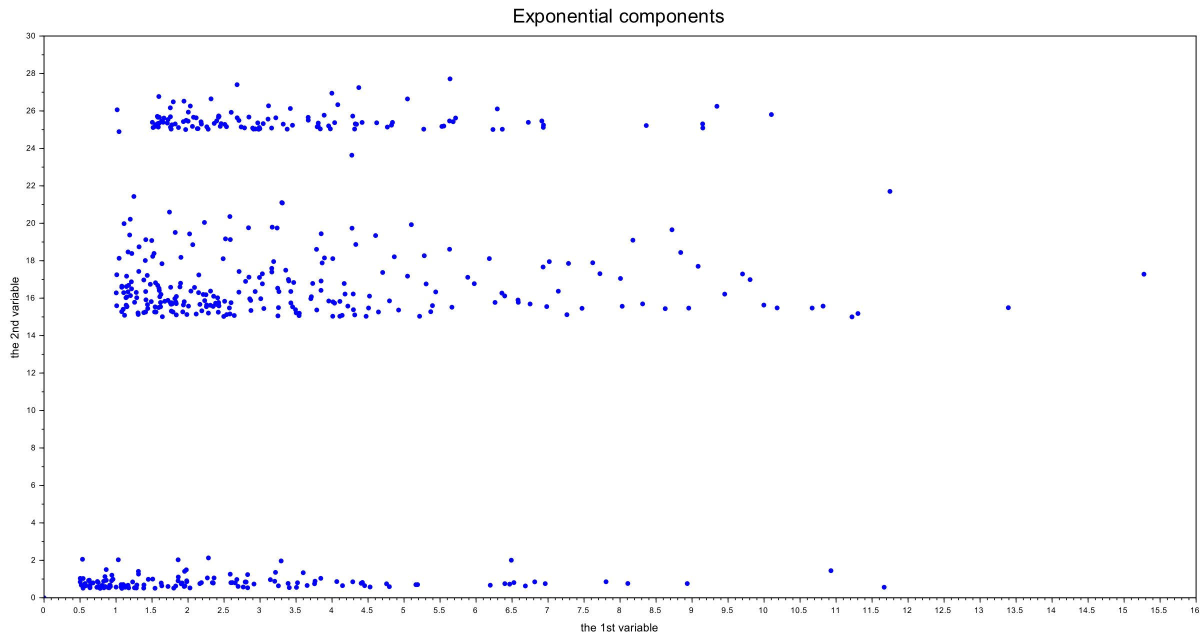


Figure 17: Simulation of exponential components: $a * 1$, $b * 0.5$

3.3.2 Results

The results of estimation and classification both using a data-dependent pointer model and a data-independent pointer model are presented in the following table. The results are valid for different values of parameters a and b. The table shows the number of incorrectly estimated samples with a data-dependent pointer model (with $z(t)$) via the number of wrong estimations using a data-independent pointer model (without $z(t)$).

Table 15: Classification results for exponential components

	b*0.01		b*0.1		b*0.2		b*0.5		b*1		b*1.5	
	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)
a*0.01	296	308	295	316	264	343	291	324	265	265	183	230
	319	310	301	297	282	283	316	328	296	288	239	320
	323	300	312	307	294	283	322	303	299	302	240	252
	329	351	320	310	303	291	323	328	303	294	279	320
	334	323	322	305	308	286	340	321	308	319	281	241
	336	316	324	269	315	299	345	339	315	276	295	286
	339	313	326	279	342	331	352	300	322	323	327	312
	341	304	333	324	344	294	353	323	323	302	344	198
	353	318	341	328	347	331	356	341	332	298	350	182
	356	317	342	303	348	298	361	298	336	288	353	203
a*0.1	281	309	303	240	186	265	83	248	36	190	25	27
	301	282	310	297	194	194	88	91	60	50	28	25
	301	302	320	299	201	211	86	196	68	50	36	30
	306	309	328	333	213	181	96	166	70	38	38	36
	309	303	336	316	260	184	104	118	79	32	44	48
	317	287	341	279	272	301	180	155	79	49	48	42
	322	277	349	328	279	160	181	57	88	71	101	77
	326	318	365	329	331	286	184	184	113	72	151	21
	339	329	365	338	342	236	193	180	165	47	158	30
	382	308	370	264	426	340	234	191	199	185	171	30
a*0.2	304	292	172	175	159	145	63	197	33	13	1	1
	311	294	214	174	169	166	67	46	41	24	2	2
	315	305	235	173	172	167	67	65	45	48	5	5
	327	355	259	297	178	172	68	63	45	152	6	5
	332	322	262	175	202	192	71	82	55	68	8	3
	333	292	305	231	214	167	71	45	59	40	16	6
	336	297	331	191	325	184	91	85	144	45	33	3
	338	271	332	172	446	293	93	74	164	31	90	14
	342	345	340	166	457	173	93	56	165	148	90	5
	348	296	341	233	468	164	94	55	182	45	91	5
a*0.5	296	280	154	175	38	37	4	5	8	25	1	1
	307	294	158	176	61	89	10	8	11	116	1	1
	323	323	169	164	67	48	12	12	27	92	3	40
	325	311	171	157	68	47	12	12	38	14	5	0
	328	307	183	170	75	52	13	13	107	37	10	169
	330	315	190	178	77	53	20	13	148	164	19	68
	340	305	313	154	77	59	21	14	156	17	19	170
	346	305	353	174	79	53	22	13	161	197	20	1
	366	313	372	169	80	46	22	18	197	35	39	146
	388	311	386	155	121	47	23	12	197	29	167	182
a*1	280	278	45	32	14	14	0	0	0	0	0	0
	300	292	52	47	16	16	0	0	0	18	0	13
	313	289	60	49	17	16	1	0	0	22	0	42
	315	257	94	71	18	13	1	1	0	50	0	46
	318	313	100	50	18	18	2	0	0	148	0	52
	331	311	120	102	25	17	2	0	1	172	0	68
	342	271	160	37	27	14	2	1	6	16	0	168
	355	272	160	98	34	16	2	1	11	9	0	169
	358	286	163	49	36	17	2	1	13	48	0	179
	377	257	164	100	36	17	2	2	21	23	1	1
a*1.5	225	278	17	17	5	4	0	0	0	12	0	0
	302	284	26	21	5	4	0	0	0	17	0	6
	304	309	27	25	6	4	0	0	0	19	0	17
	307	255	28	20	6	5	0	0	0	26	0	22
	320	262	30	24	7	4	0	0	0	33	0	24
	321	301	31	28	7	5	0	0	0	35	0	41
	347	214	31	28	7	7	0	0	0	38	0	44
	350	292	40	35	8	7	0	0	0	69	0	54
	358	218	43	29	10	7	3	0	73	17	0	171
	392	228	45	35	10	8	4	0	210	154	0	172

There are some examples illustrating simulation and classification results and the difference in the results while using a pointer model with $z(t)$ and without $z(t)$.

The first example is with parameter a multiplied by 0.01 and b multiplied by 0.01.

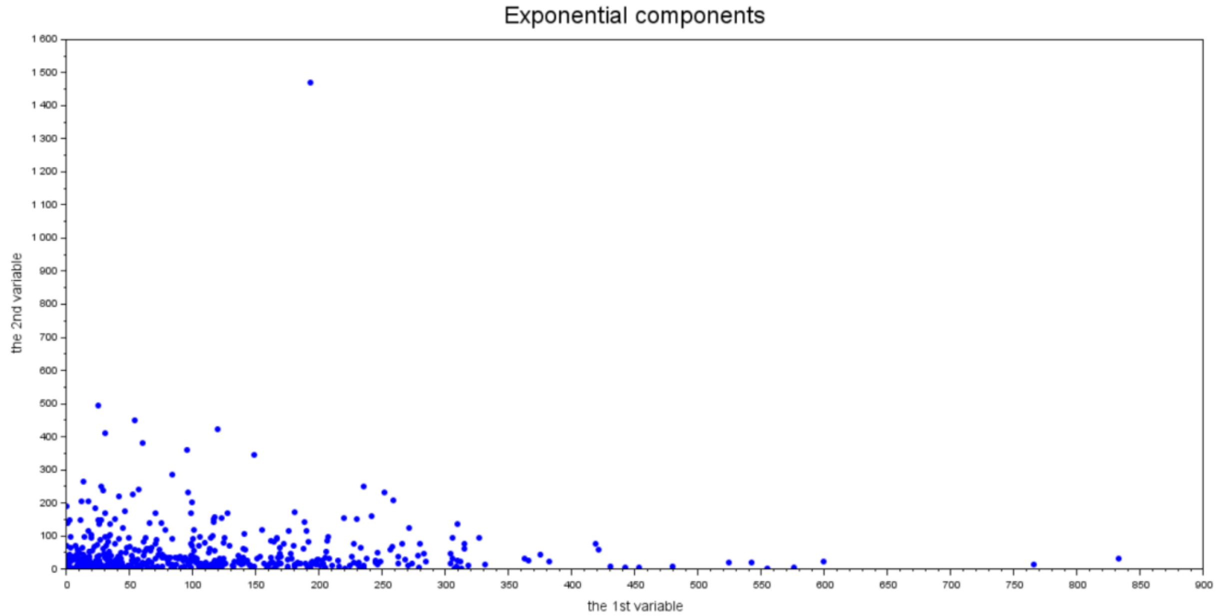
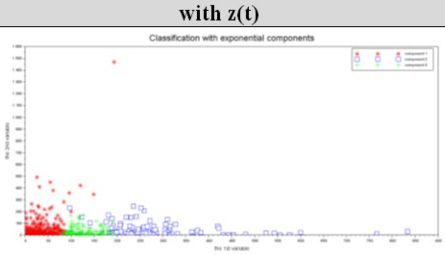
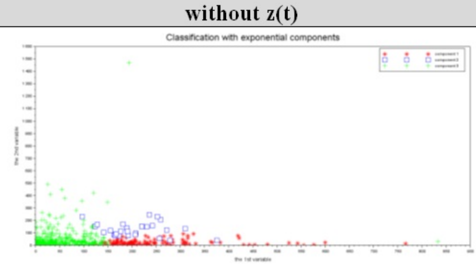
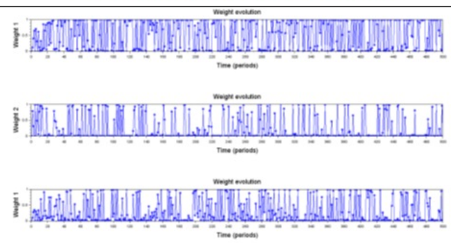
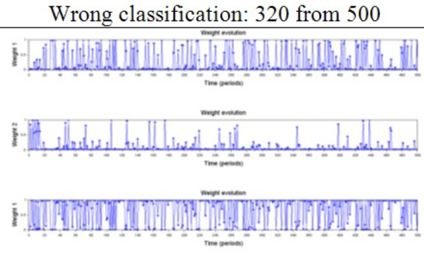
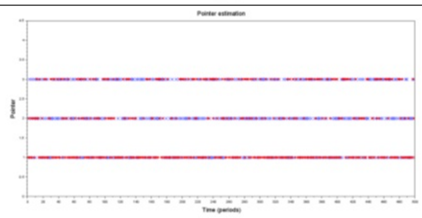
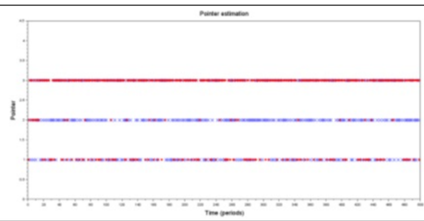


Figure 18: Simulation of exponential components: $a * 0.01$, $b * 0.01$

The results of estimation and classification for the data-dependent and data-independent pointer models are listed in the following table.

Table 16: Classification results: $a * 0.01$, $b * 0.01$

with z(t)	without z(t)
 Wrong classification: 349 from 500	 Wrong classification: 320 from 500
	
	

The next example is for parameter a multiplied by 0.5 and b multiplied by 0.01.

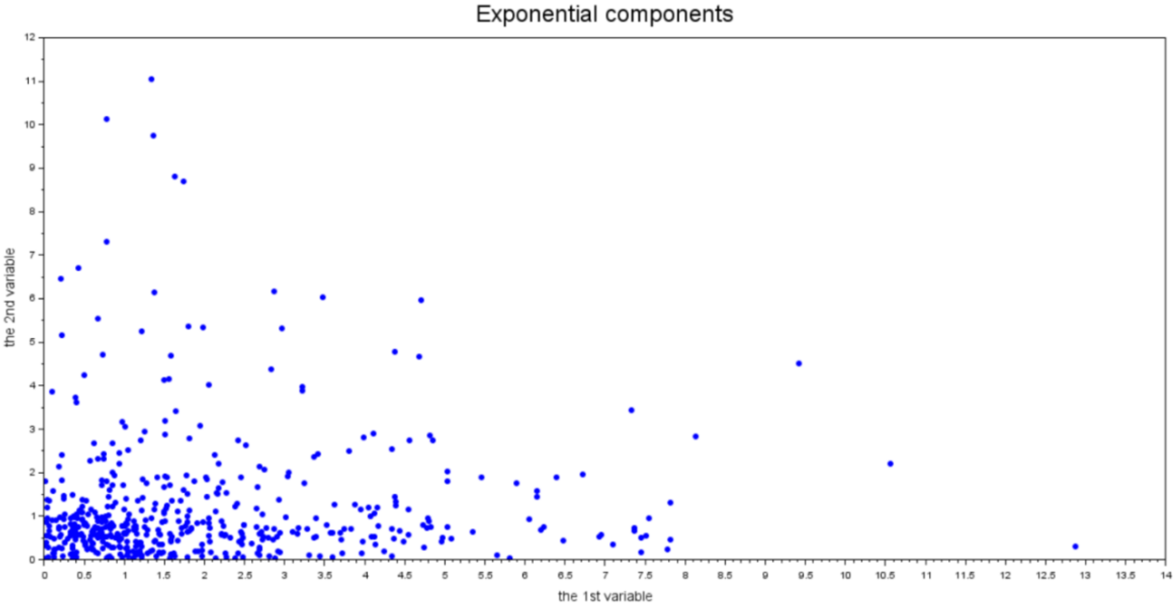
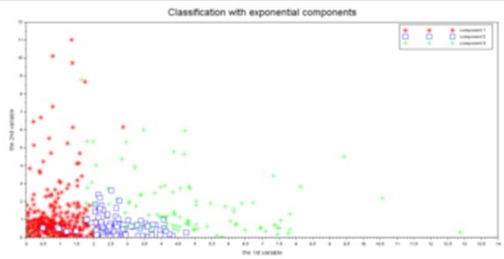
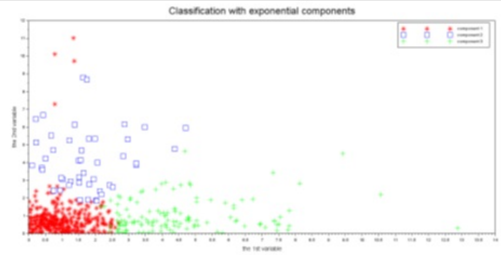
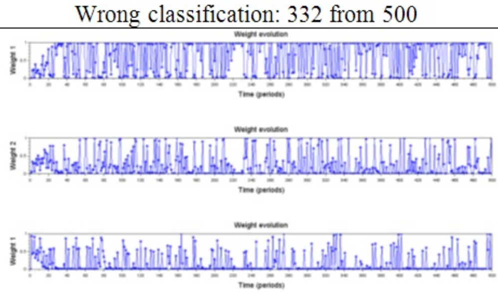
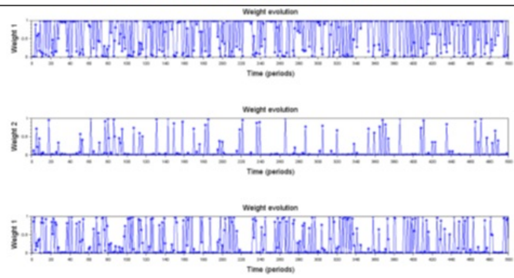
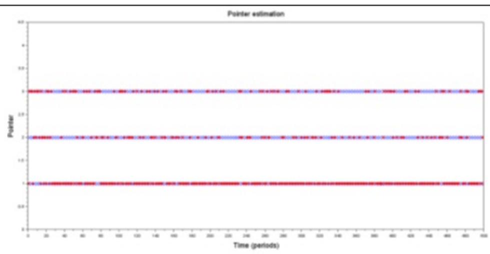
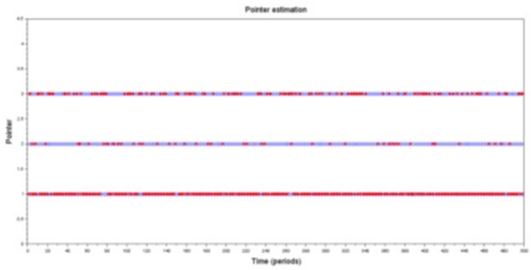


Figure 19: Simulation of exponential components: $a * 0.5$, $b * 0.01$

Table 17: Classification results: $a * 0.5$, $b * 0.01$

with z(t)	without z(t)
 Wrong classification: 332 from 500	 Wrong classification: 283 from 500
	
	

The simulation results for parameter a multiplied by 1 and parameter b multiplied by 0.01 look as follows.

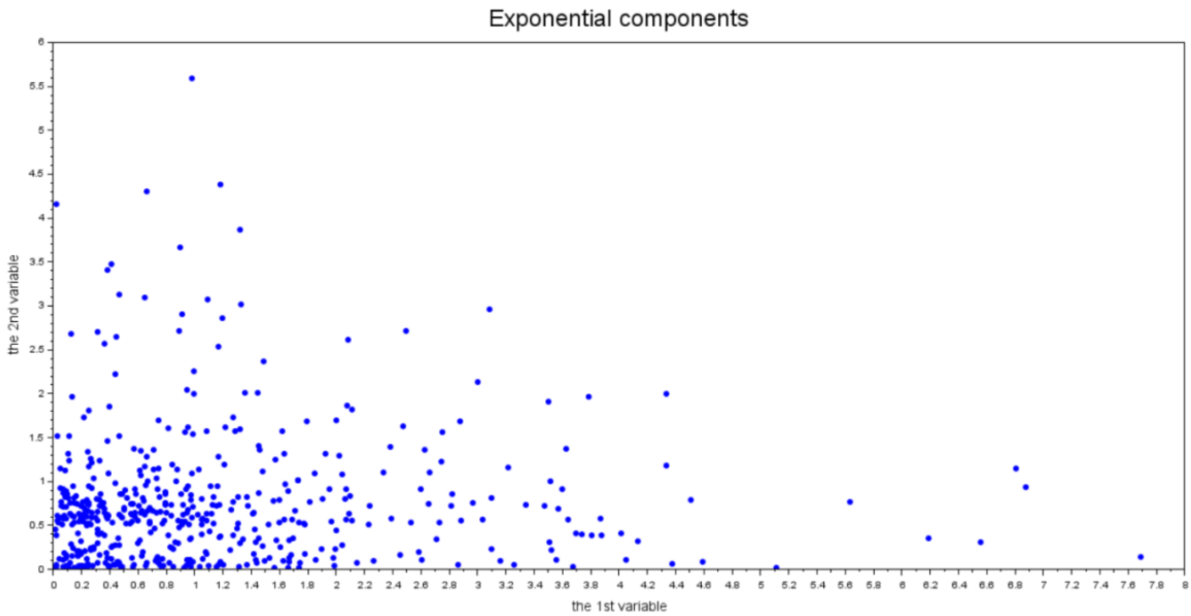
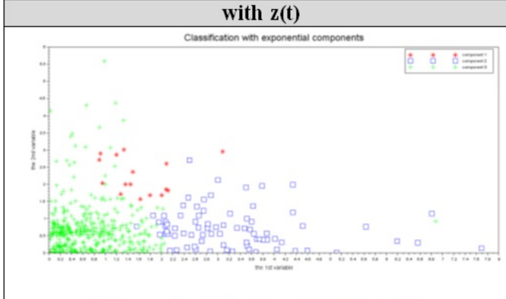
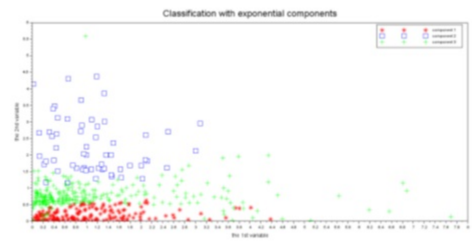
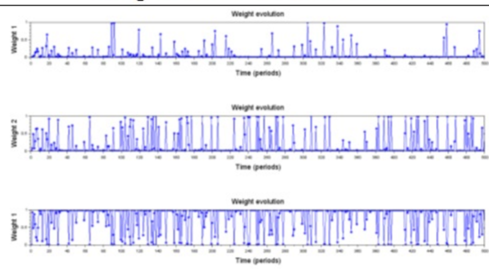
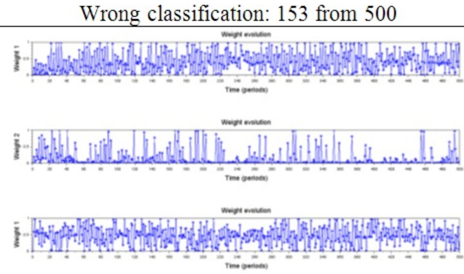
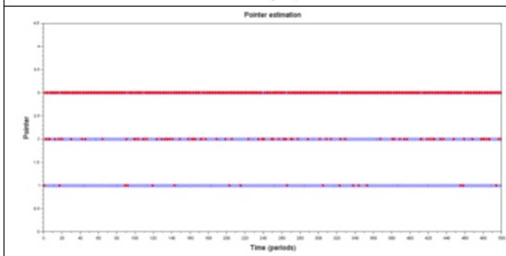
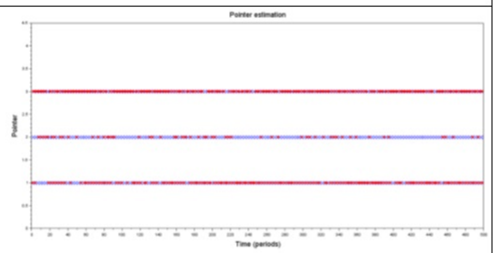


Figure 20: Simulation of exponential components: $a * 1$, $b * 0.01$

Table 18: Classification results: $a * 1$, $b * 0.01$

with z(t)	without z(t)
 Wrong classification: 320 from 500	 Wrong classification: 153 from 500
	
	

The figures and table below show the results for parameters a and b multiplied by 0.5.

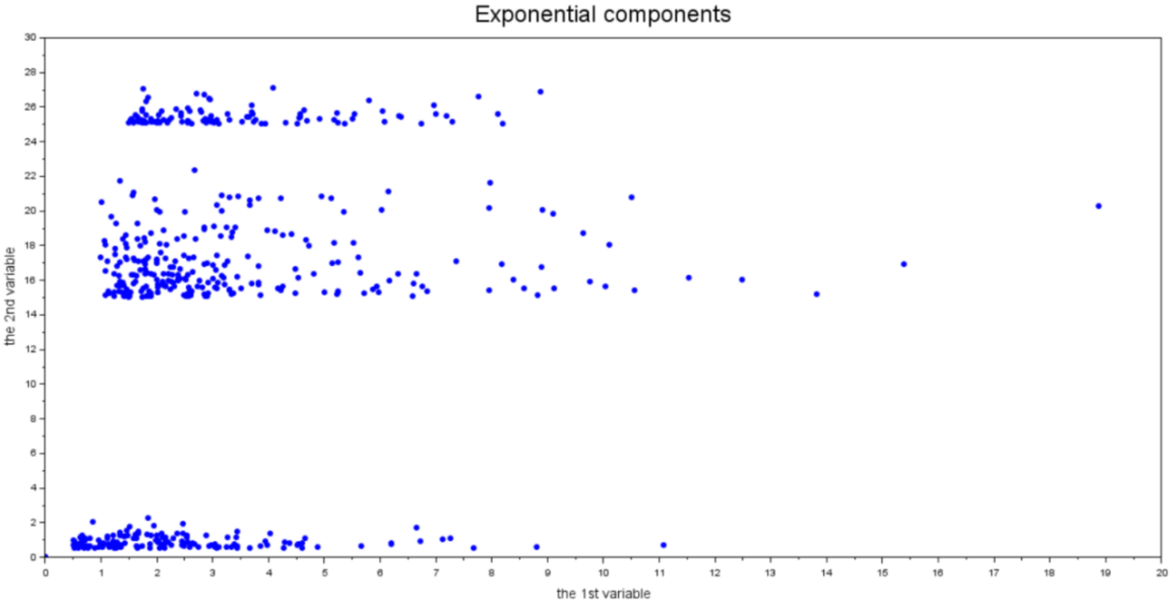
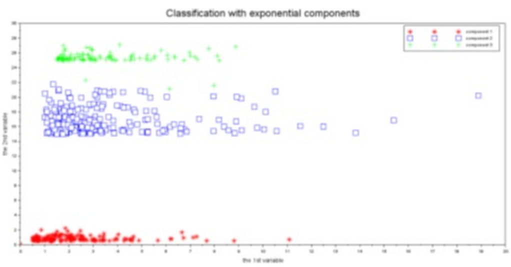
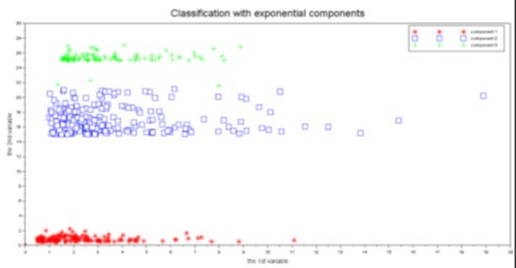
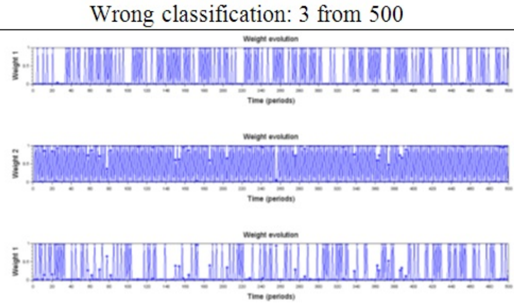
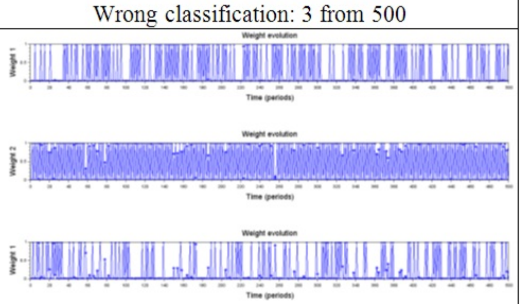
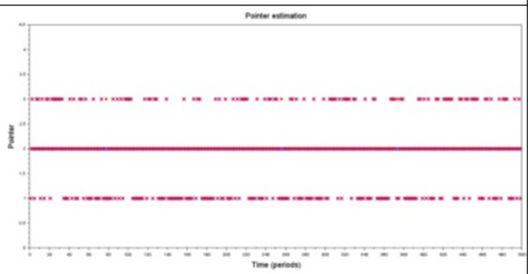
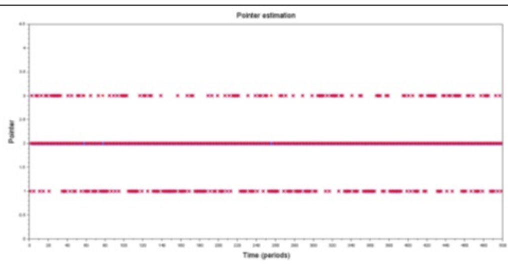


Figure 21: Simulation of exponential components: $a * 0.5$, $b * 0.5$

Table 19: Classification results: $a * 0.5$, $b * 0.5$

with z(t)	without z(t)
 Wrong classification: 3 from 500	 Wrong classification: 3 from 500
	
	

Simulation and estimation results for parameter a multiplied by 1 and parameter b multiplied by 0.5 are shown below.

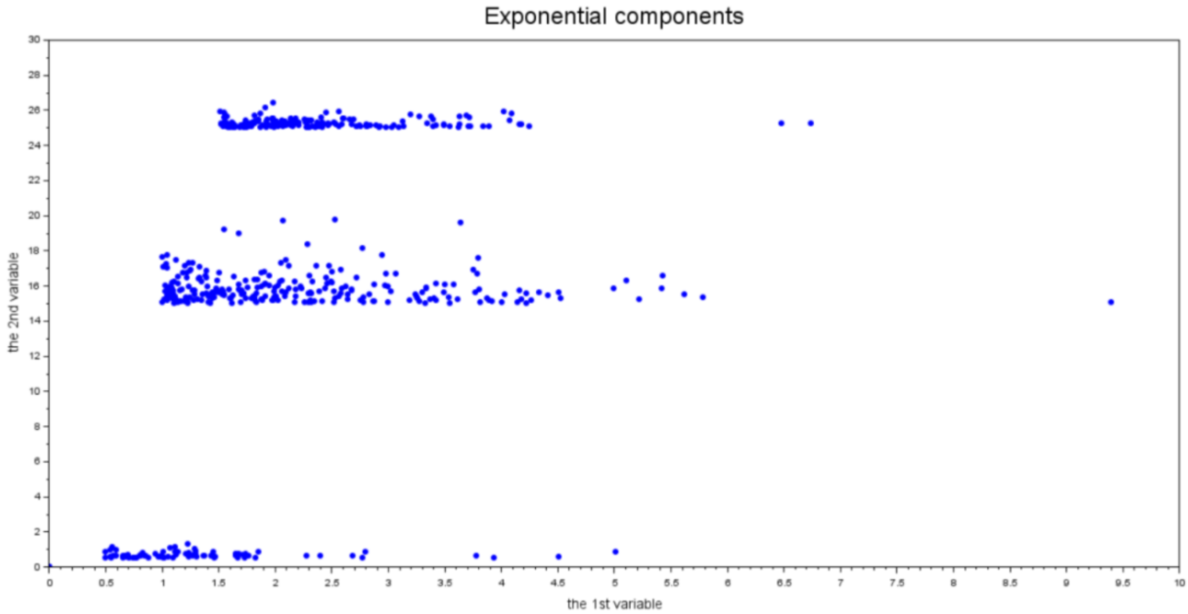
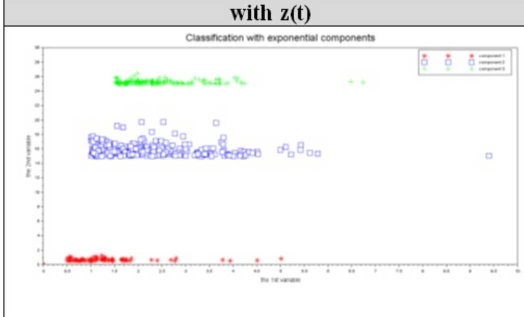
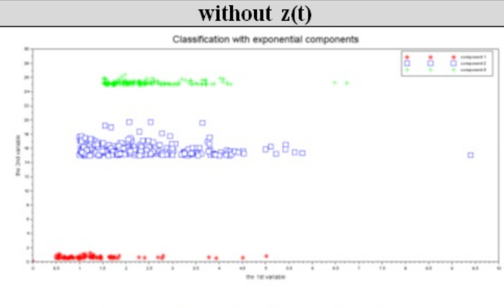
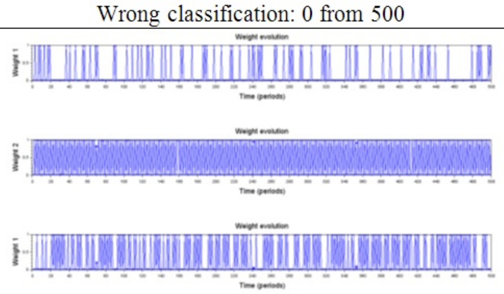
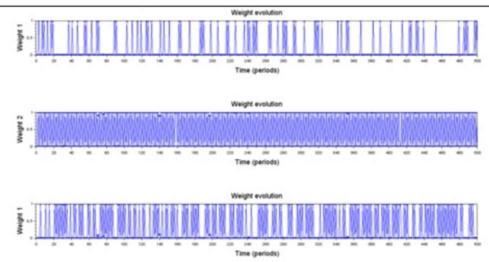
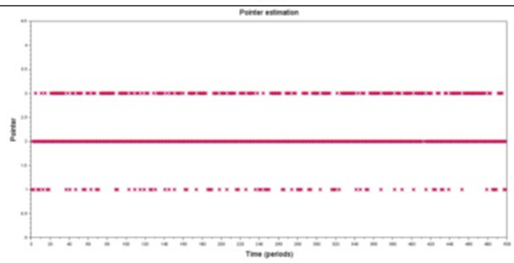
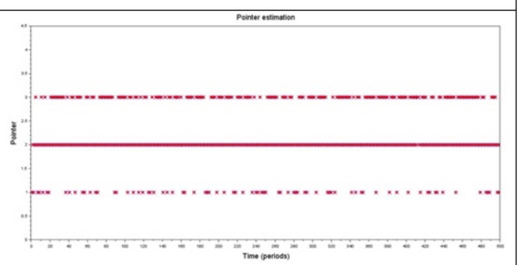


Figure 22: Simulation of exponential components: $a * 1$, $b * 0.5$

Table 20: Classification results: $a * 1$, $b * 0.5$

with z(t)	without z(t)
 Wrong classification: 0 from 500	 Wrong classification: 0 from 500
	
	

The last two examples are for parameter a multiplied by 1 and 1.5 and parameter b multiplied by 1.5.

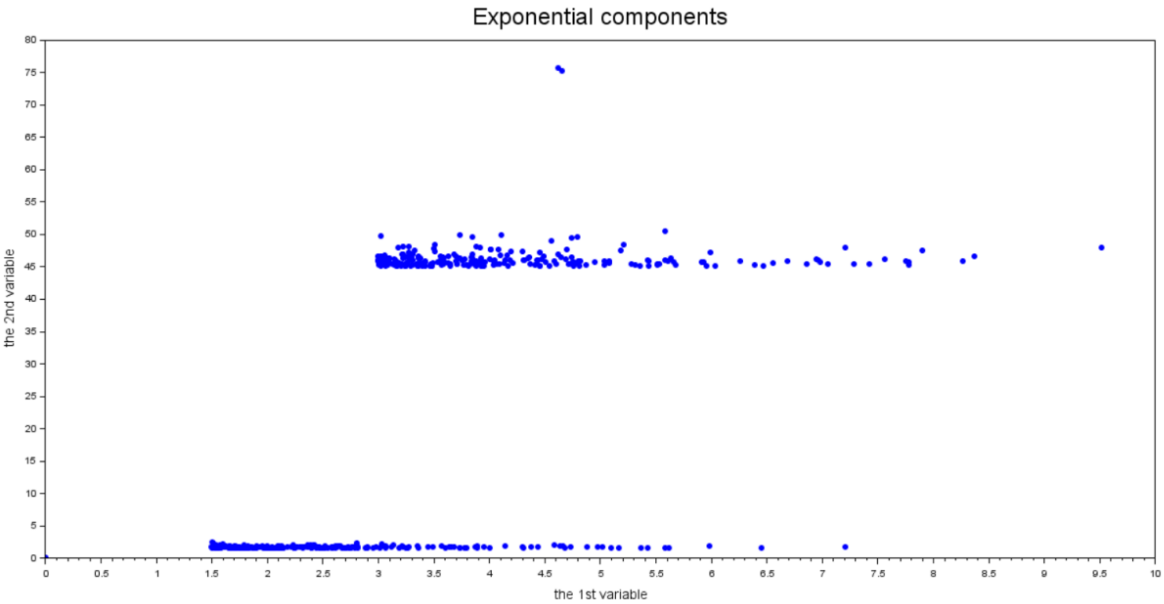
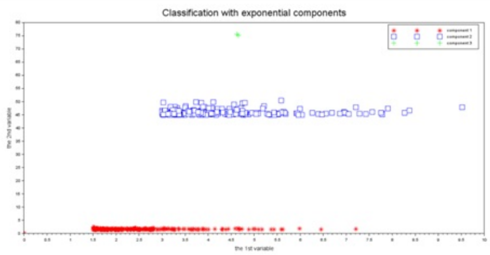
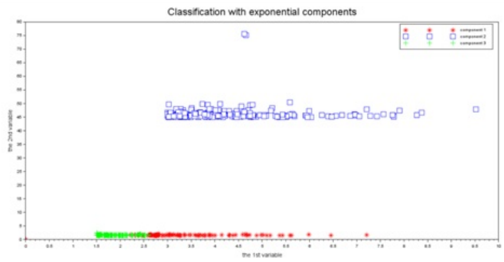
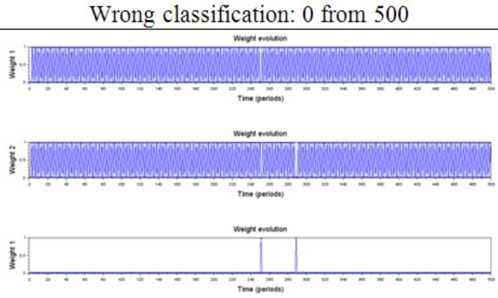
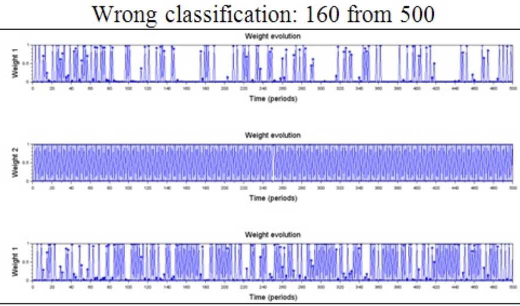
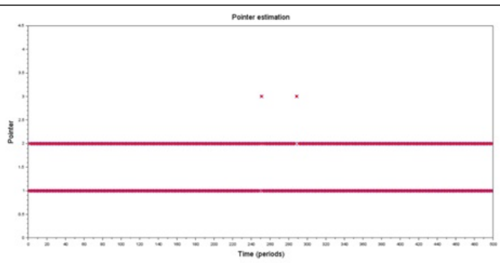
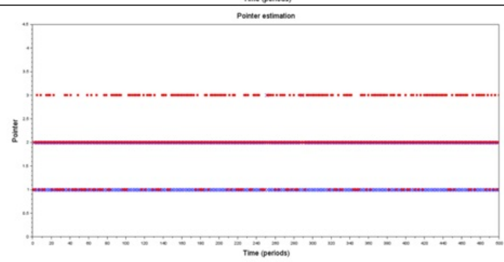


Figure 23: Simulation of exponential components: $a * 1$, $b * 1.5$

Table 21: Classification results: $a * 1$, $b * 1.5$

with z(t)	without z(t)
 Wrong classification: 0 from 500	 Wrong classification: 160 from 500
	
	

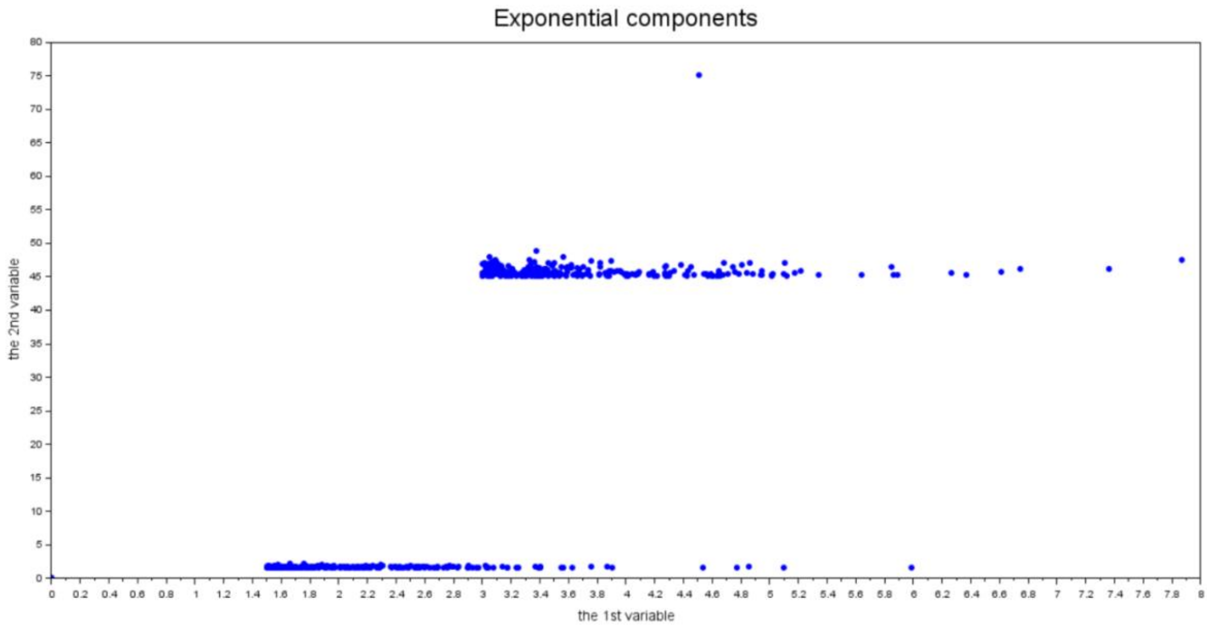
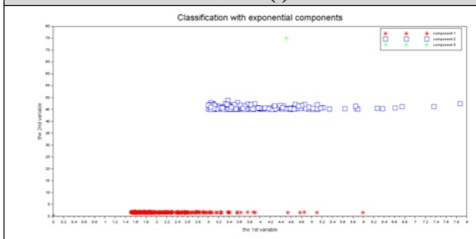
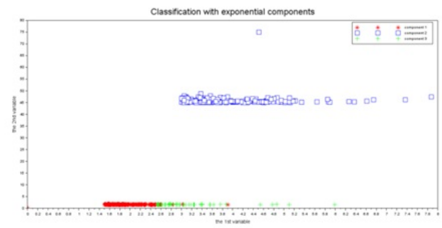
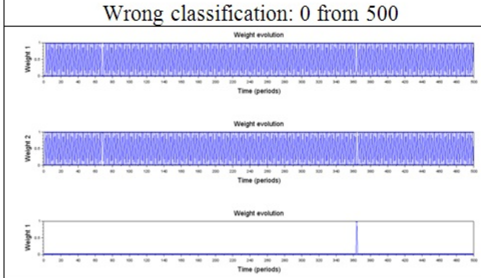
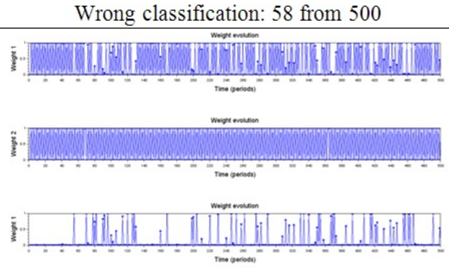
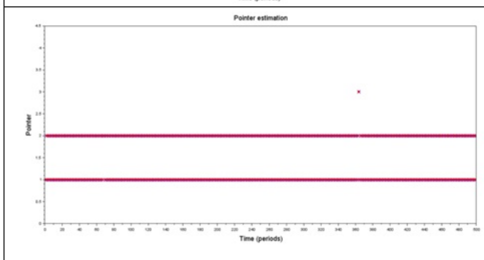
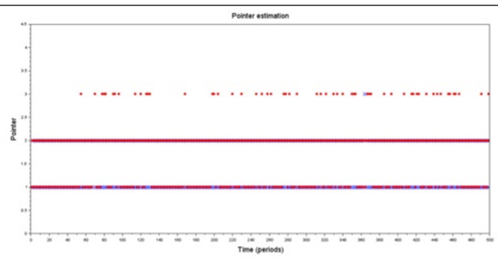


Figure 24: Simulation of exponential components: $a * 1.5$, $b * 1.5$

Table 22: Classification results: $a * 1.5$, $b * 1.5$

with z(t)	without z(t)
 Wrong classification: 0 from 500	 Wrong classification: 58 from 500
	
	

In the next section the results presented here are discussed in details.

3.3.3 Discussion

The results obtained for exponential components are quite contradictory. On the one hand the variant with a data-independent pointer model presents better results almost in all cases of estimation and classification compared with a data-dependent pointer model (see table 15 and 23). For example, for parameter a and b multiplied by 0.01, i.e. the components are very closely located, the number of misclassifications using a data-dependent pointer model varies from 296 to 356 samples (59.2 % to 71.2 %) from total number of 500 samples, while the number of incorrectly classified samples using a data-independent pointer model is in the range from 300 to 351 samples (60 % to 70.20 %) from total number of 500 samples. For a multiplied by 1 and b multiplied by 0.5 the number of wrong classifications with a pointer model with $z(t)$ is 1.4 samples (0.28 %) in average, while the number of incorrectly estimated samples using a pointer model without $z(t)$ is 0.6 (0.12 %) samples in average.

The case when parameter a is multiplied by 0.1 and b is multiplied by 0.5 is rather exclusion than the rule. Here the number of misclassification using a data-dependent pointer model is lower (142.9 samples in average) than using a data-independent pointer model (158.6 samples in average).

As it was already noted, the greater values of the parameters are, i.e. the greater distance between components are, the more accurate the classification is. But from the certain point the data-independent pointer model fails, while a data-dependent pointer model presents very good results. This is the case when a and b are multiplied by 1 and 1.5 (see tables and figures presented in the previous section). Thus, for instance, for parameters a and b multiplied by 1.5 the number of misclassifications for a data-dependent pointer model is 0, while the number of incorrectly classified samples using a pointer model without $z(t)$ is from 0 to 172 samples (0 % to 34.4 %), i.e. 55.1 samples or 11.02 % in average. In a data-independent pointer model the samples tend to be allocated more or less equal between all three groups, thus, failing to estimate the real allocation of data.

The average and standard deviation values for different values of parameters are listed in table 23. Red cells in the table represents a greater number of wrong classifications with a data-dependent pointer model in comparison with a data-independent pointer model. The better classification results with $z(t)$ compared to the model without $z(t)$ are coloured in green.

Table 23: Classification results: average and standard deviation

	b*0.01		b*0.1		b*0.2		b*0.5		b*1		b*1.5	
	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)
a*0.01	332.60±17.38	316±14.10	321.60±15.54	303.80±18.46	314.70±29.83	303.90±22.40	335.90±22.25	320.50±15.37	309.90±20.80	295.50±17.72	289.10±56.39	254.40±52.46
a*0.1	318.40±27.46	302.40±16.22	338.70±23.67	302.30±32.84	270.40±77.65	235.80±59.89	142.90±56.65	158.60±55.93	95.70±50.17	78.40±58.84	80±59.25	36.60±16.30
a*0.2	328.60±14.27	306.90±26.06	279.10±59.68	198.70±42.32	279±131.54	182.30±40.83	77.80±13.08	76.80±44.41	93.30±61.70	61.40±48.98	34.20±39.80	4.90±3.57
a*0.5	334.90±26.96	306.40±11.97	244.90±97.91	167.20±9.13	74.30±20.65	53.10±13.88	15.90±6.52	12±3.46	105±76.98	72.60±66.22	28.40±50.11	77.80±80
a*1	328.90±29.46	282.60±19.58	111.80±48.58	63.50±27.14	24.10±8.71	15.80±1.62	1.40±0.84	0.60±0.70	5.20±7.44	50.60±59.97	0.10±0.32	73.80±71.29
a*1.5	322.60±44.68	264.10±34.59	31.80±8.57	26.20±6.05	7.10±1.79	5.50±1.58	0.70±1.49	0	28.30±67.84	42±42.63	0	55.10±63.58

Besides all mentioned above another problem in the process of estimation and classification of exponential components with a data-dependent pointer model was in the fact, that it estimated the samples belonging to the same group correctly, but referred them to the wrong group. The example of this error is obvious from the following figures.

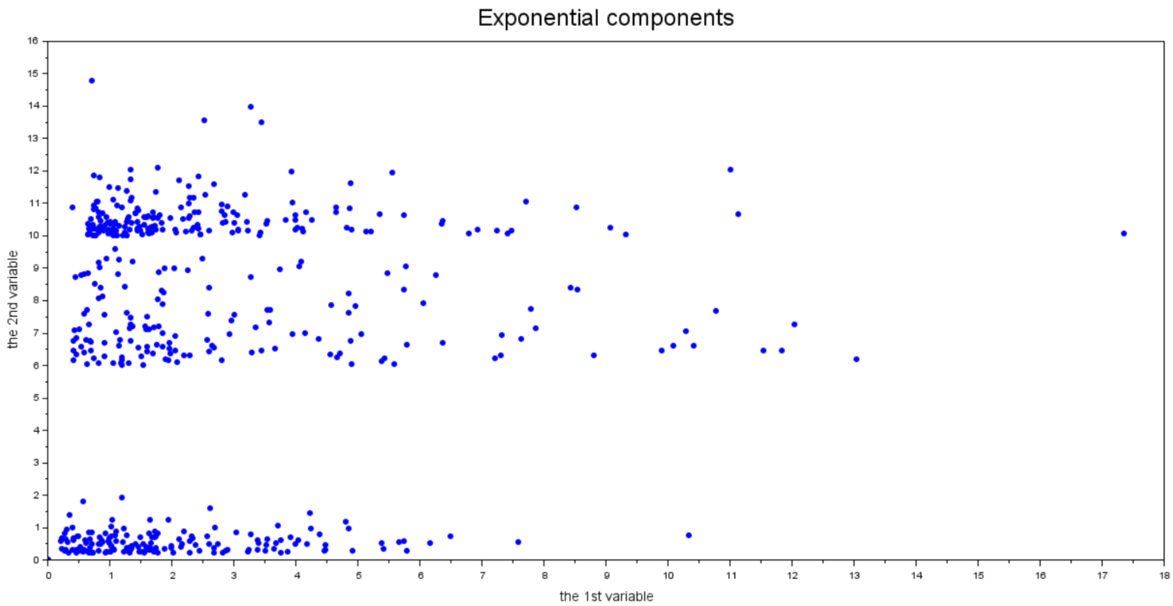
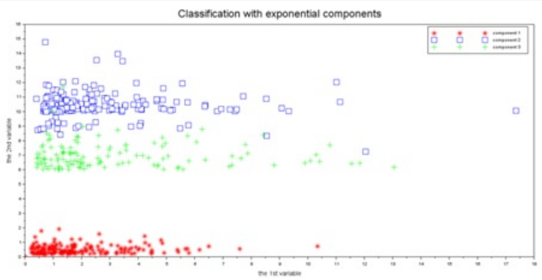
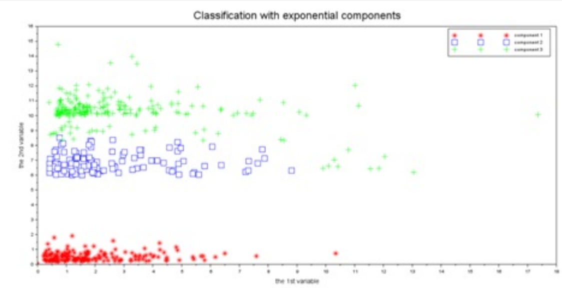


Figure 25: Simulation of exponential components: $a * 0.5$, $b * 0.2$

Table 24: Classification results: $a * 0.5$, $b * 0.2$

	
A data-dependent pointer model: 293 incorrectly classified samples from total number of 500 samples	**A data-independent pointer model: 54 incorrectly classified samples from total number of 500 samples**

From the figures above it is clear that classification results are very similar. The problem is that the second (blue) and the third (green) components in the figure on the left should be interchanged. This mistake results in a great number of incorrectly classified samples with a data-dependent model compared with a data-independent model: 293 samples via 54 samples resp.

This problem was particularly obvious for the following multiplication values of parameters a and b:

Table 25: Examples of parameter values when the errors in classification occur

$a * 0.1$	$b * 1$
$a * 0.2$	$b * 0.5$
$a * 0.5$	$b * 0.2, b * 0.5, b * 1, b * 1.5$
$a * 1$	$b * 0.1, b * 0.2$
$a * 1.5$	$b * 0.1, b * 0.2$

Because of the possibility of such an error one should be very careful not to include such results in final evaluation of the performance of the algorithm with the data-dependent and data-independent pointer models.

3.4 State-space components

3.4.1 Data simulation

The parameters for three components and parameters for a pointer model are defined in the first step of simulation. The number of simulated data is 150, including data for a component model $y(t)$, $x(t)$ and $u(t)$, data for a pointer model $c(t)$, discrete values for a pointer model $z(t)$ and weights α.

During simulation the values of noise for $x(t)$ and $y(t)$ were changed. r_v is a parameter for changing noise ratio in $y(t)$ and r_w is a parameter for changing noise ratio in $x(t)$. The values of r_w and r_v during simulation were 0.001, 0.1, 1, 5 and 10. Thus, there were 25 different combinations and for each combination ten experiments were performed. The total number of experiments is 250.

Several examples of simulation with different noise ratio are shown in the following figures.

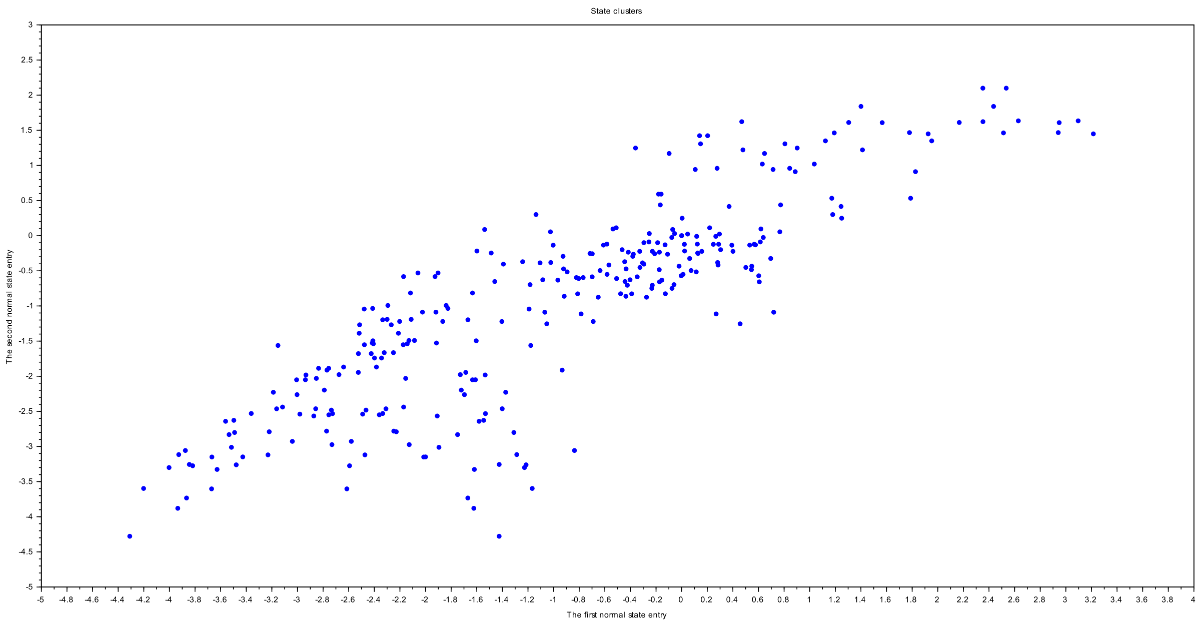


Figure 26: Simulation of state-space components for $r_w = 0.001$, $r_v = 0.001$

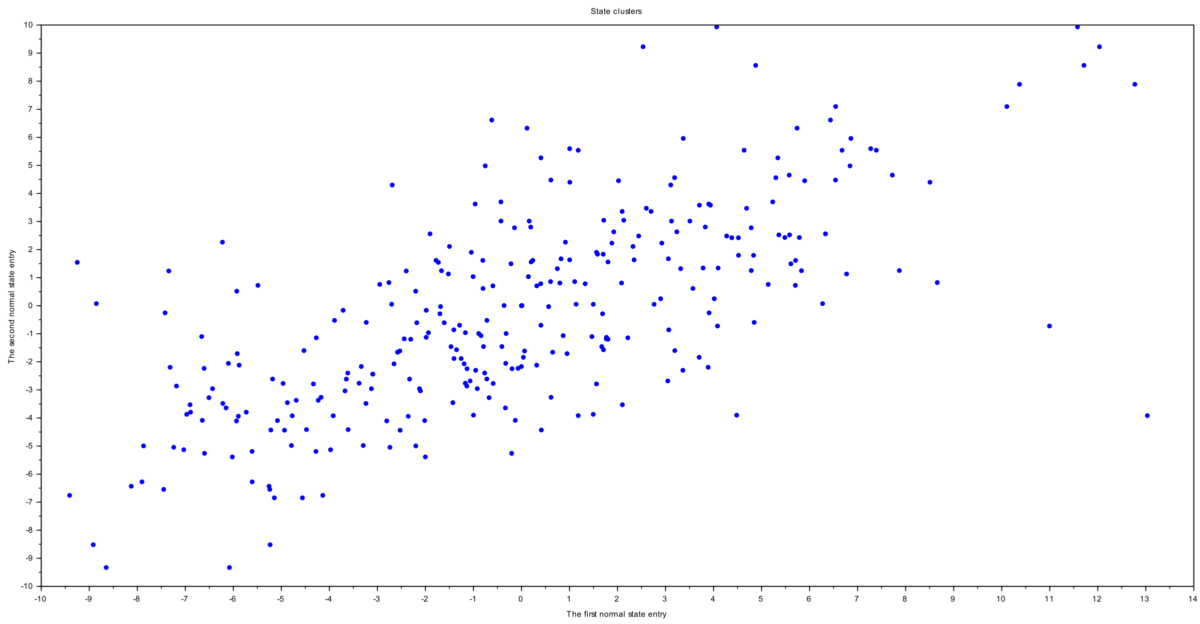


Figure 27: Simulation of state-space components for $r_w = 10$, $r_v = 0.001$

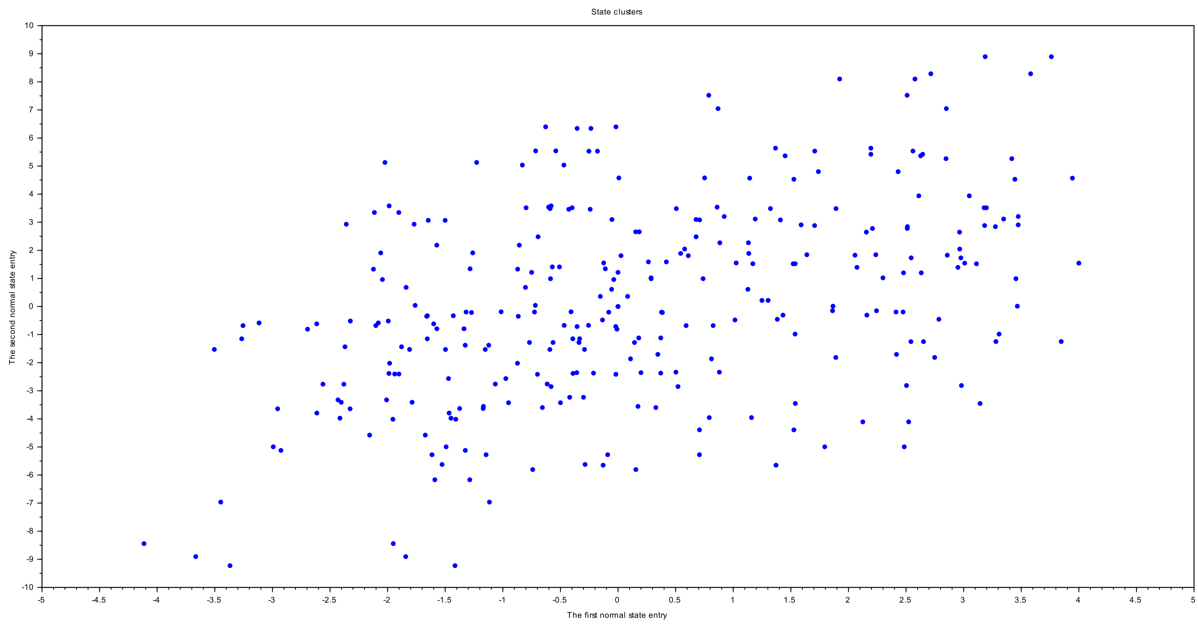


Figure 28: Simulation of state-space components for $r_w = 0.001$, $r_v = 10$

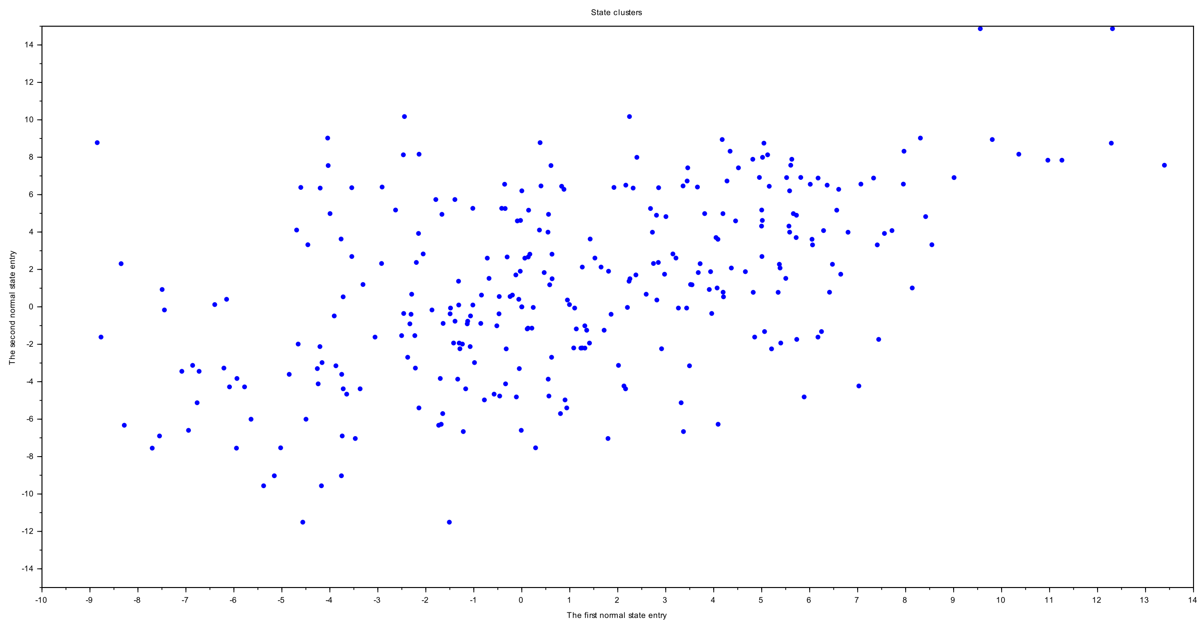


Figure 29: Simulation of state-space components for $r_w = 10$, $r_v = 10$

3.4.2 Results

The results of estimation and classification with a data-dependent pointer model (with $z(t)$) and a data-independent pointer model (without $z(t)$) are listed in the following table. The results are presented for various values of noise both in $y(t)$ and $x(t)$.

Table 26: Classification results for state-space components

	rv=0.001		rv=0.1		rv=1		rv=5		rv=10	
	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)	with z(t)	without z(t)
rw=0.001	0	0	1	1	31	32	56	56	68	67
	0	0	1	2	34	34	59	60	76	76
	0	0	3	3	34	35	60	61	77	76
	0	0	3	4	37	37	63	63	77	77
	0	0	4	3	37	40	66	66	79	80
	0	0	4	4	39	39	66	67	80	79
	0	0	5	6	39	40	67	69	80	80
	0	1	6	6	40	41	70	69	81	81
	0	1	12	14	41	41	71	70	83	83
	2	1	13	13	42	42	76	76	89	88
rw=0.1	1	1	5	5	34	34	59	59	67	67
	2	4	7	8	37	38	63	64	69	71
	3	5	9	9	41	41	66	67	74	73
	4	3	10	10	42	42	66	68	75	75
	4	4	10	10	42	42	68	68	77	77
	4	5	10	13	42	44	70	70	80	79
	5	7	11	11	46	46	71	72	80	80
	7	4	12	12	47	46	72	71	82	82
	7	5	13	15	49	49	73	73	84	84
	8	7	15	15	50	51	76	76	89	88
rw=1	28	29	29	29	47	47	59	59	72	72
	29	27	31	31	48	48	60	58	76	76
	29	29	32	35	49	49	67	67	77	77
	30	30	36	37	51	50	69	70	77	77
	30	30	37	36	53	54	69	70	77	77
	35	33	37	37	56	55	70	70	78	78
	35	35	38	40	56	56	71	71	78	78
	35	35	42	43	56	57	72	70	78	78
	36	37	45	45	57	57	75	74	81	81
	38	38	47	48	59	59	78	78	85	88
rw=5	59	59	58	58	60	60	68	69	70	70
	60	62	61	61	61	61	70	70	73	73
	61	59	61	61	67	68	71	73	76	77
	61	61	63	63	68	69	73	73	77	76
	63	63	64	65	70	71	76	76	77	79
	66	66	65	65	71	72	76	77	78	78
	69	70	67	68	72	72	77	77	80	79
	70	70	68	69	73	72	81	81	80	81
	71	71	69	67	73	73	84	84	87	87
	73	73	72	73	76	76	85	85	92	91
rw=10	69	67	71	70	67	67	73	72	76	77
	72	72	71	71	72	72	73	73	78	78
	72	72	73	73	74	74	76	76	81	81
	75	76	75	75	76	75	77	76	81	81
	77	77	76	75	77	77	79	78	84	84
	77	77	76	75	77	77	79	82	84	84
	79	78	77	78	78	77	82	80	84	84
	84	85	78	78	78	79	83	83	86	86
	84	84	79	78	81	81	87	86	87	85
	85	87	79	80	82	81	89	89	90	89

There are several examples of simulation and estimation results for various values of noise illustrated in figures below.

The first example is for noise values r_v=0.001 and r_w=0.001.

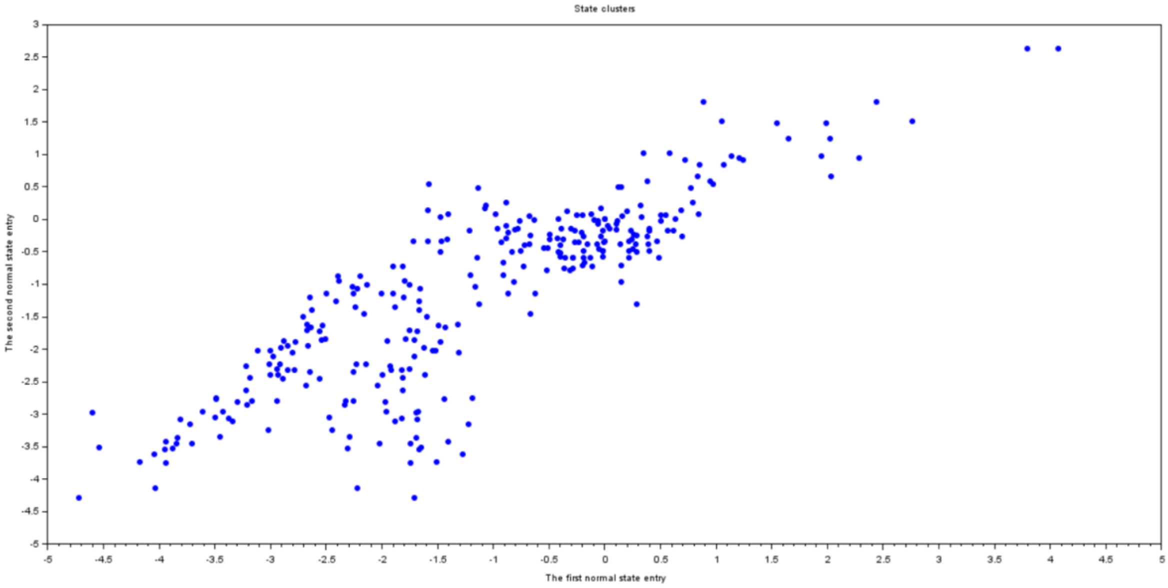
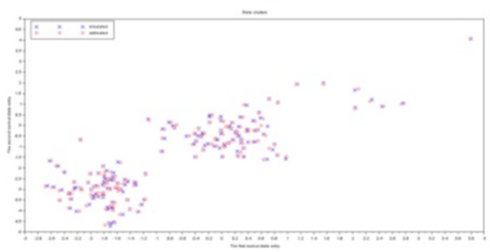
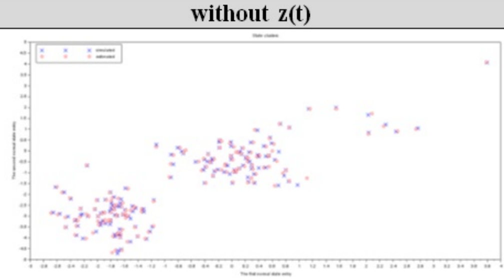
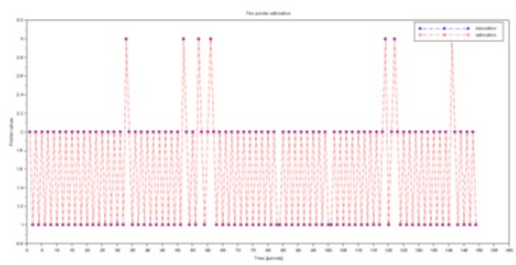
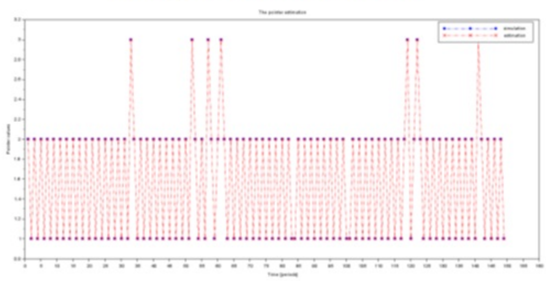
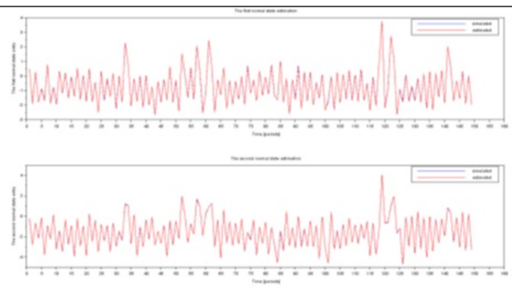
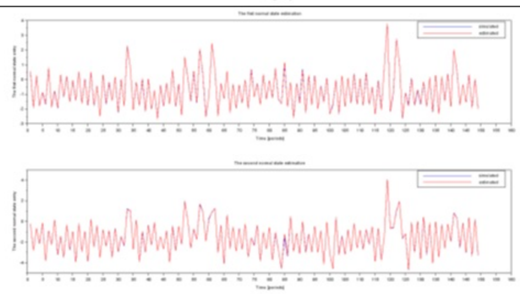


Figure 30: Simulation of state-space components for $r_w = 0.001$, $r_v = 0.001$

Table 27: Classification results forfor $r_w = 0.001$, $r_v = 0.001$

with z(t)	without z(t)
 Wrong classifications: 0 of 150	 Wrong classifications: 0 of 150
	
	

The next example is for noise values r_w=0.001, r_v=1.

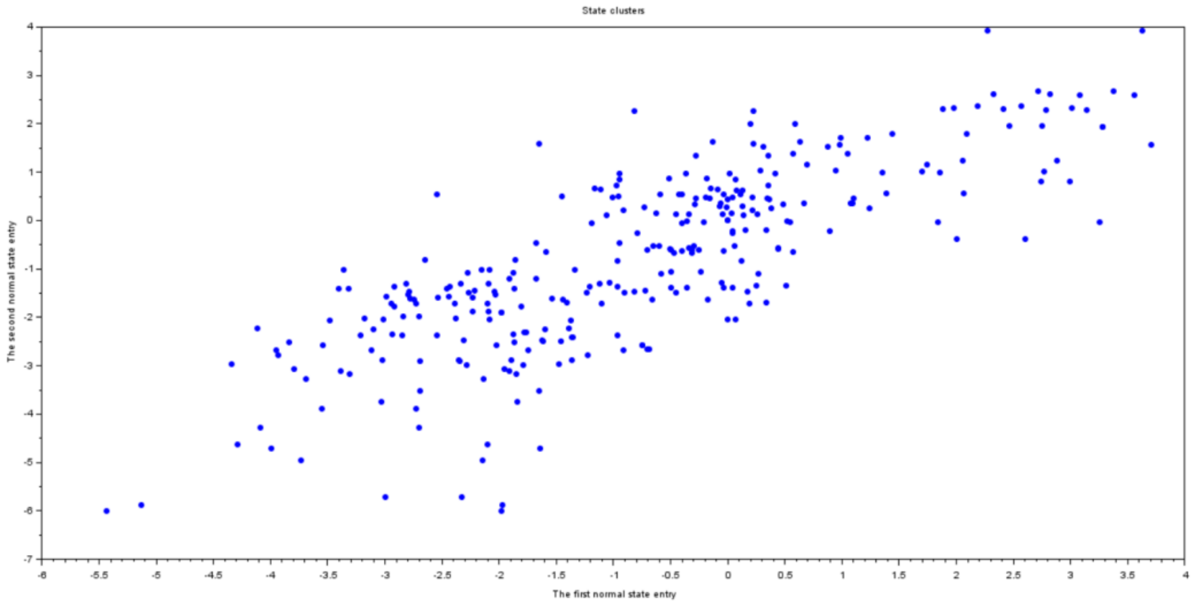
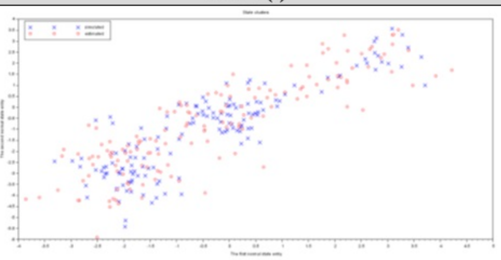
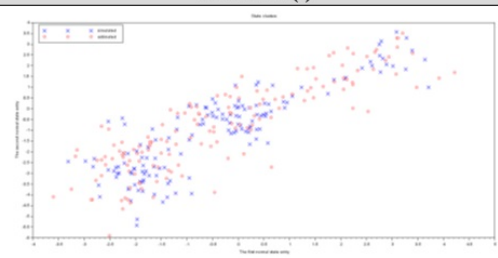
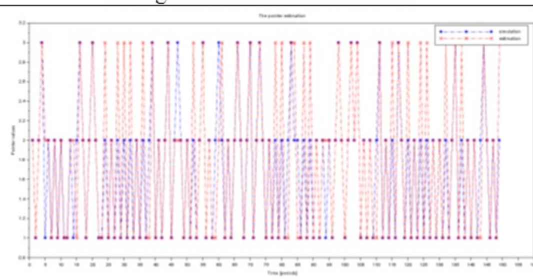
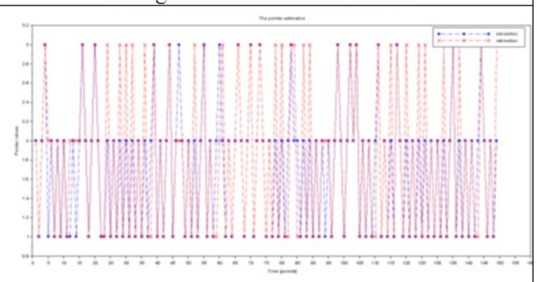
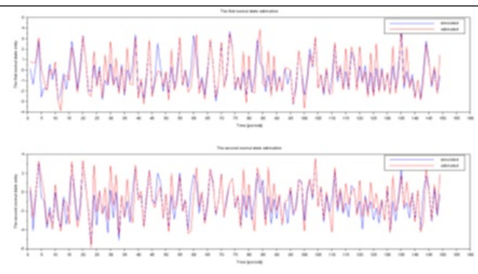
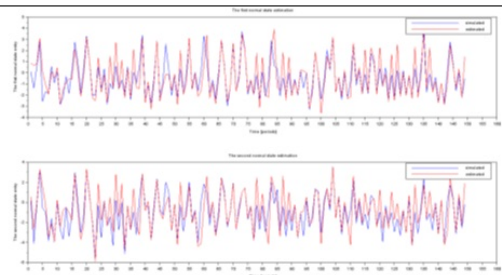


Figure 31: Simulation of state-space components for $r_w = 0.001$, $r_v = 1$

Table 28: Classification results forfor $r_w = 0.001$, $r_v = 1$

with z(t)	**without z(t)**
 Wrong classifications: 31 of 150	 Wrong classifications: 31 of 150
	
	

The example for noise values r_w=0.001, r_v=10 are presented below.

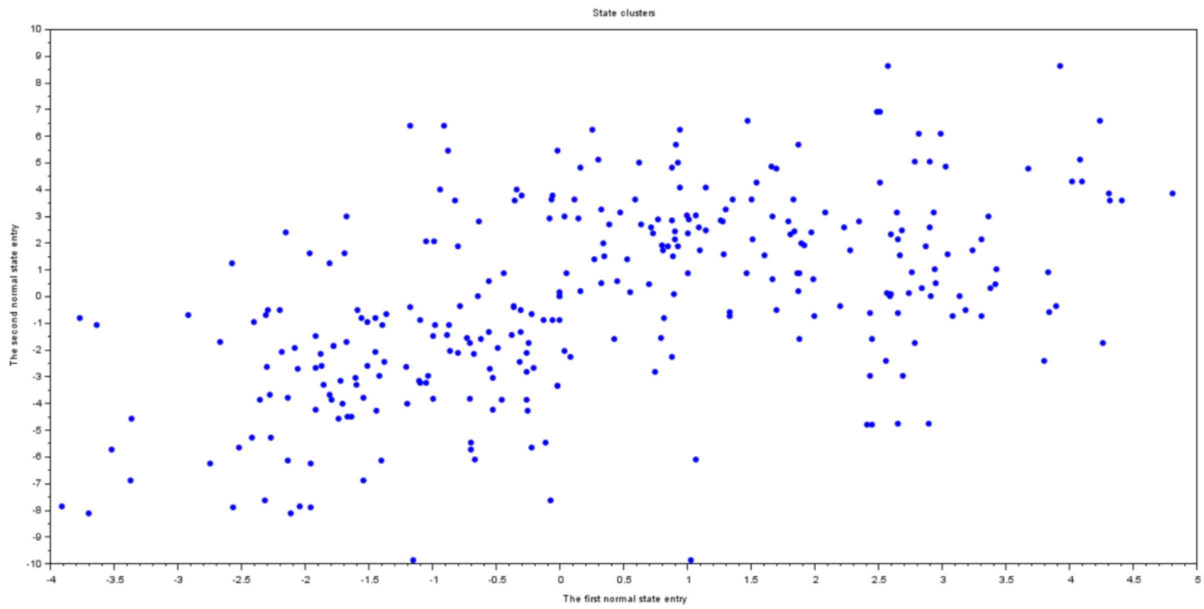
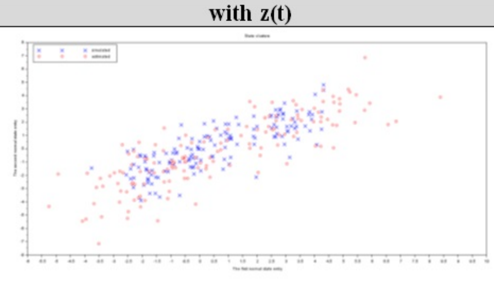
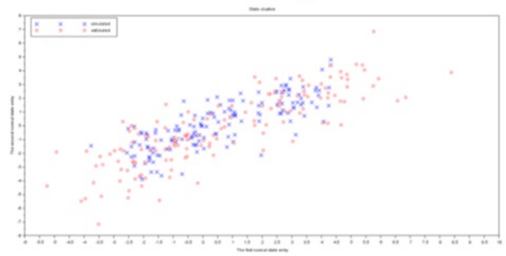
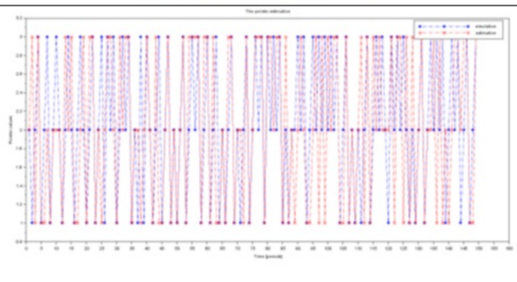
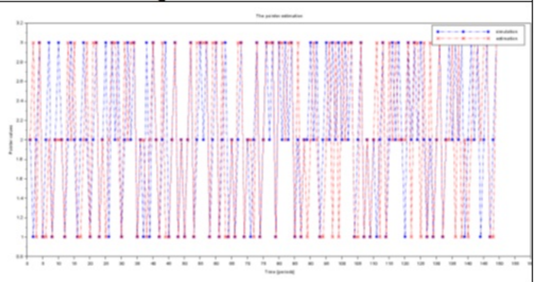
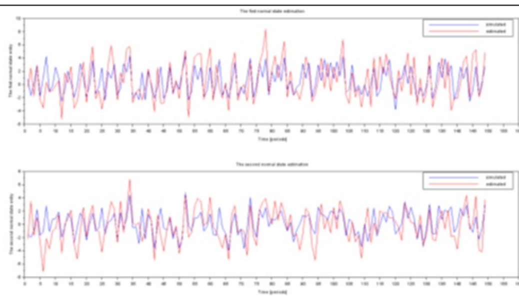
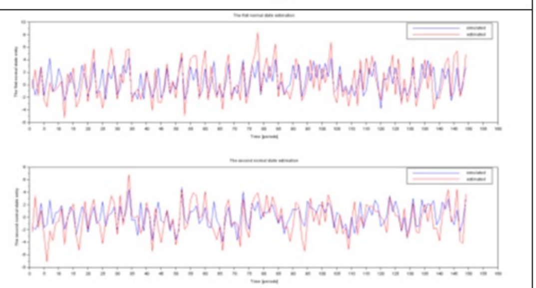


Figure 32: Simulation of state-space components for $r_w = 0.001$, $r_v = 10$

Table 29: Classification results forfor $r_w = 0.001$, $r_v = 10$

with z(t)	without z(t)
 Wrong classifications: 75 of 150	 Wrong classifications: 75 of 150
	
	

The experiment with noise values r_w=10, r_v=0.001 is illustrated in the following figures.

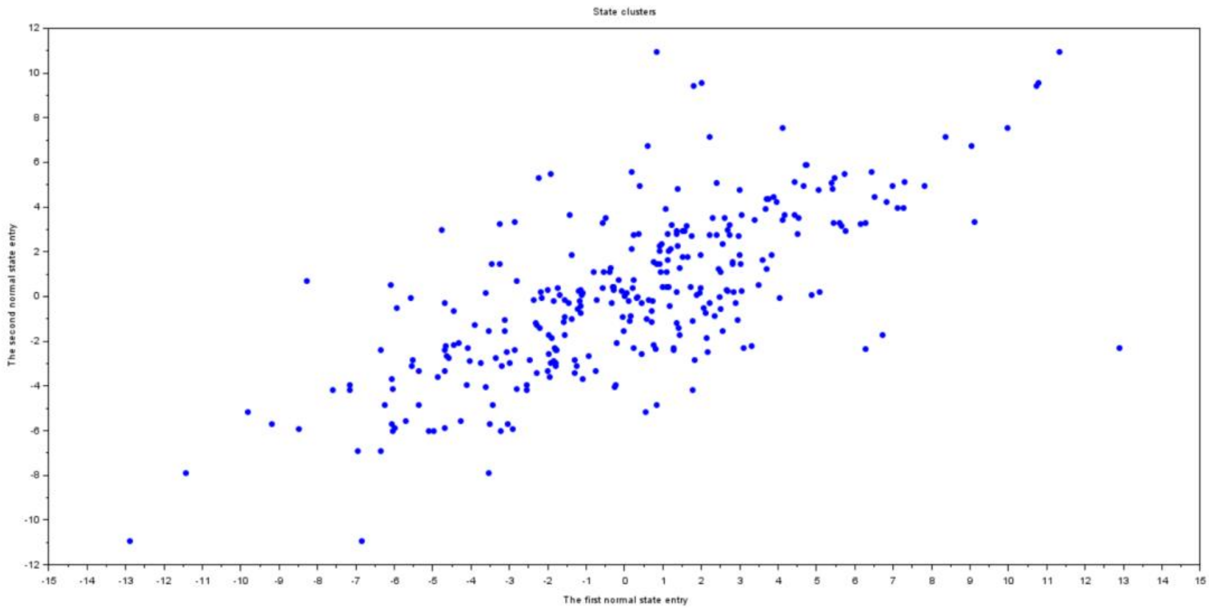
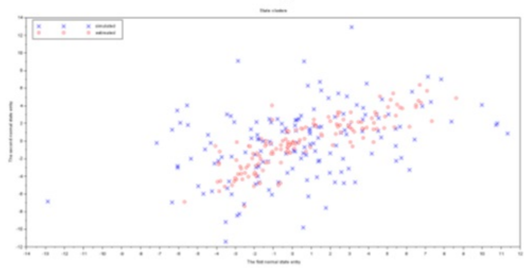
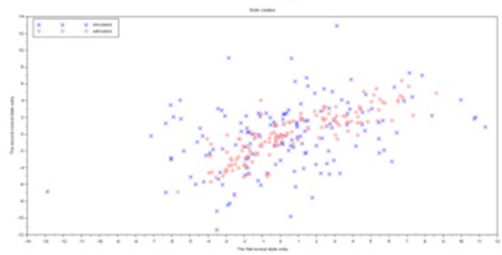
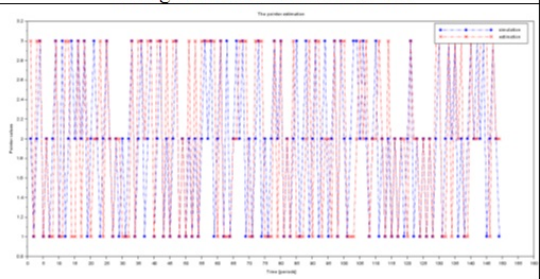
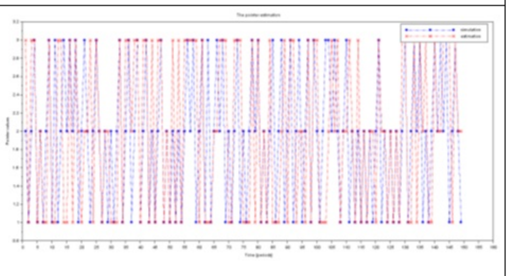
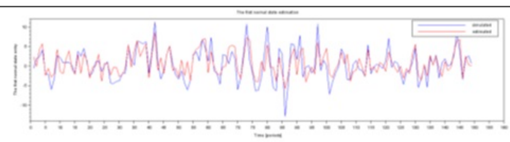
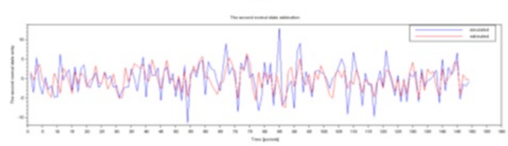
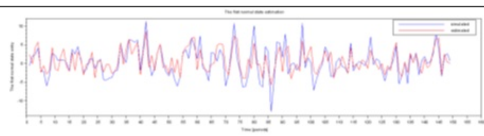
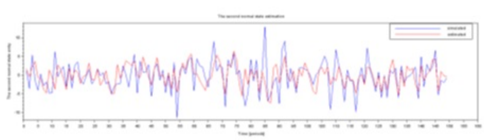


Figure 33: Simulation of state-space components for $r_w = 10$, $r_v = 0.001$

Table 30: Classification results forfor $r_w = 10$, $r_v = 0.001$

with z(t)	without z(t)
 Wrong classifications: 74 of 150	 Wrong classifications: 74 of 150
	
 	 

Two last examples are for noise values r_w=10 and r_v=1, and r_w=10 and r_v=10.

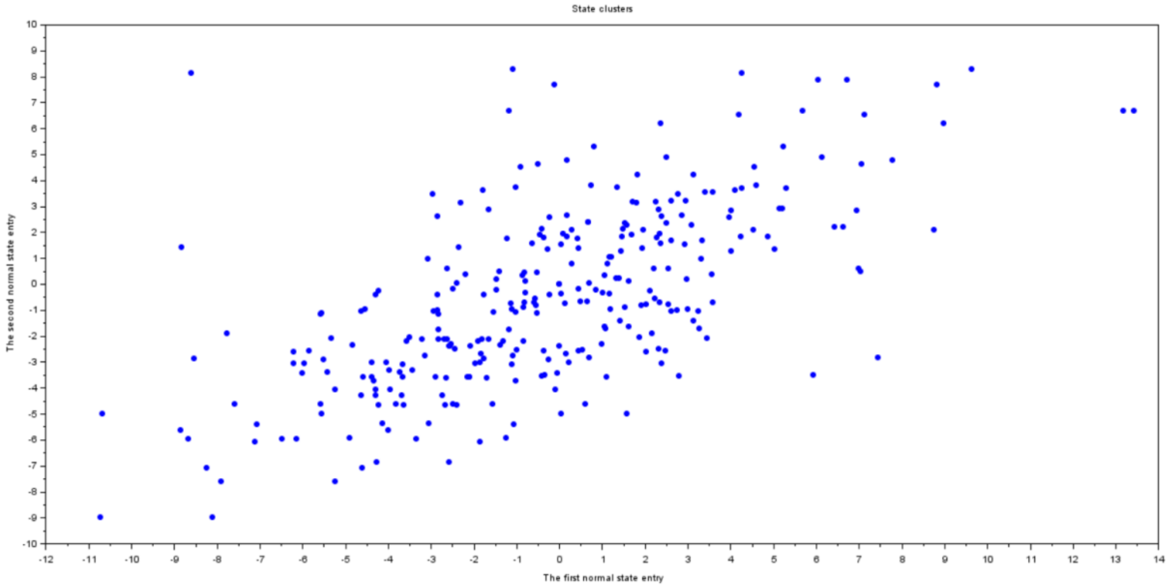
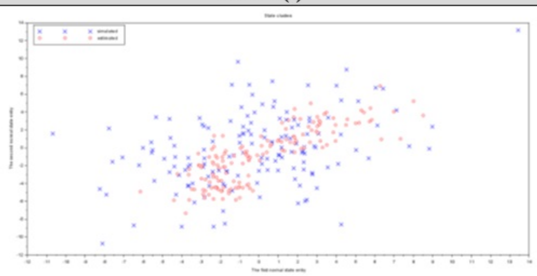
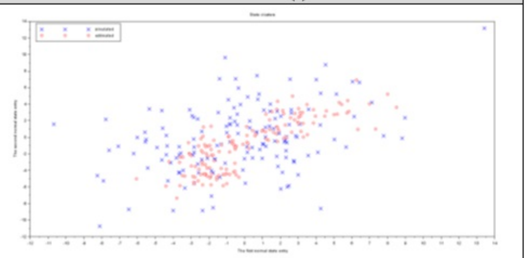
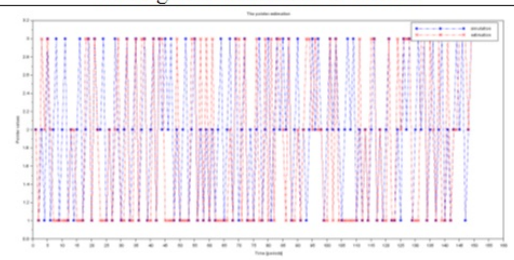
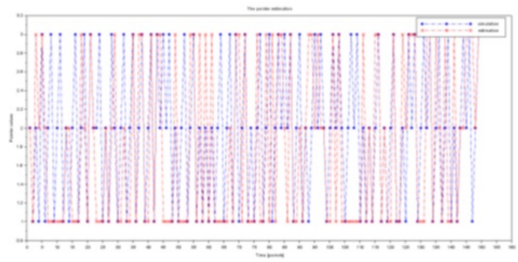
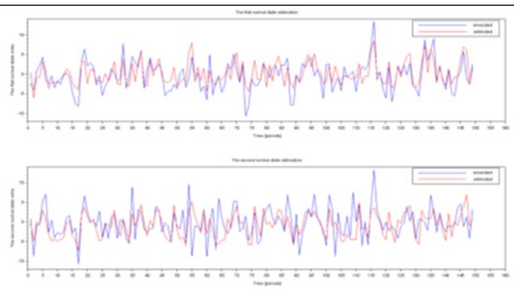
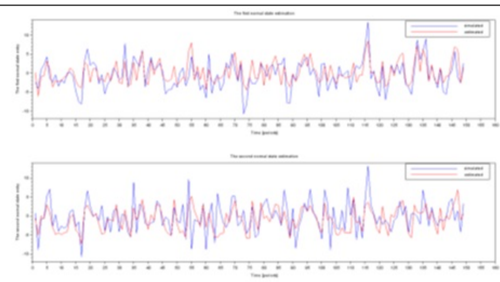


Figure 34: Simulation of state-space components for $r_w = 10$, $r_v = 1$

Table 31: Classification results forfor $r_w = 10$, $r_v = 1$

with z(t)	without z(t)
 Wrong classifications: 77 of 150	 Wrong classifications: 79 of 150
	
	

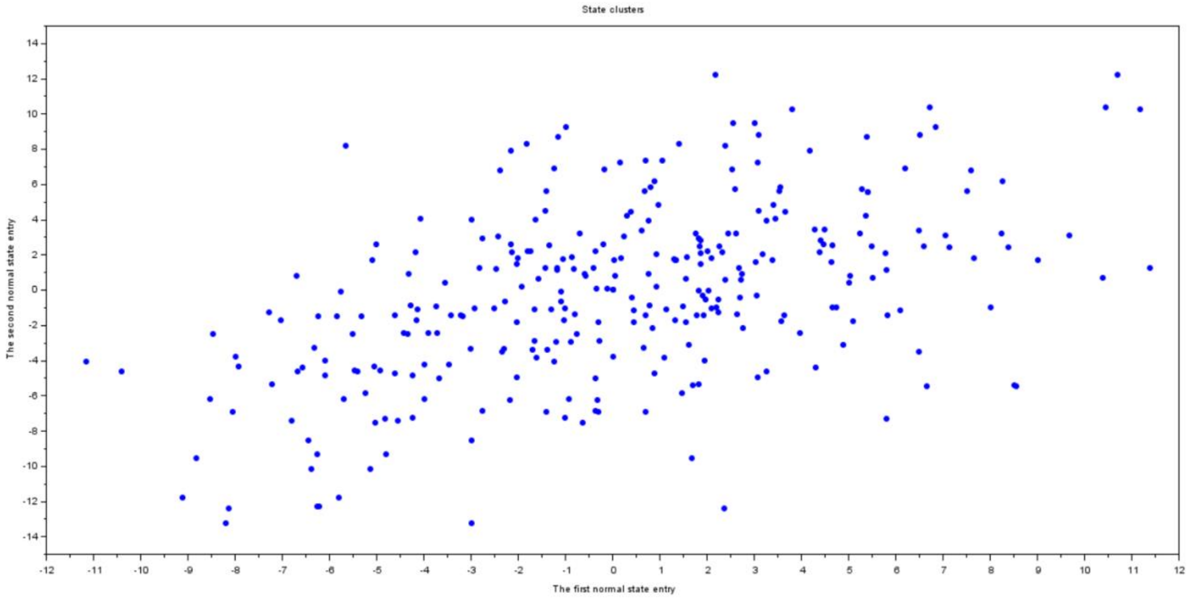
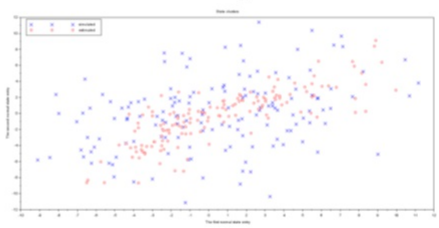
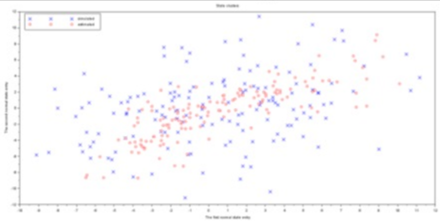
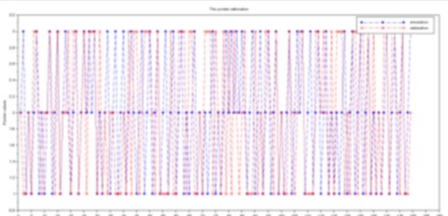
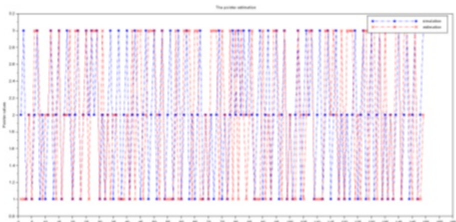
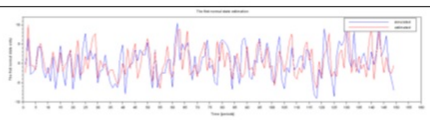
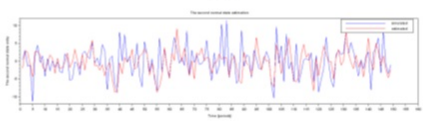
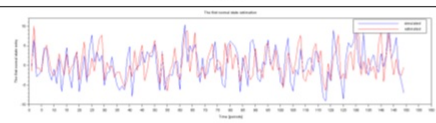
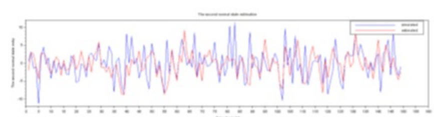


Figure 35: Simulation of state-space components for $r_w = 10$, $r_v = 10$

Table 32: Classification results forfor $r_w = 10$, $r_v = 10$

with z(t)	without z(t)
 Wrong classifications: 89 of 150	 Wrong classifications: 90 of 150
	
 	 

The results presented in this section are discussed in Section 3.4.3.

3.4.3 Discussion

According to the results obtained during experiments and presented in the previous section estimation with a data-dependent model (with $z(t)$) is slightly better than with a data-independent pointer model (without $z(t)$). Thus, for r_v=0.1 and r_w=0.001 the number of misclassifications using a pointer model with $z(t)$ is from 1 to 13 samples (0.67 % to 8.67 %) from total number of 150 samples, while the number of incorrectly estimated samples using a pointer model without $z(t)$ is from 1 to 14 samples (0.67 % to 9.33 %) from total number of 150 samples. For r_v=10 and r_w=5 again a data-dependent model has a slightly better results: 79 incorrectly classified samples (52.67 %) in average via 79.1 wrong classifications (52.73 %) in average with a data-independent pointer model (see table 26 and 33 for more details). The difference in the number of misclassifications is actually negligible.

It is also obvious from the obtained outputs that the greater value of noise is the worse the results of estimation are. For the greater values of noise the data-independent model sometimes gives more precise classification. Especially it is valid for r_w=10 in combination with r_v=0.1 and higher (see table 33). Thus, for r_w=10 and r_v=10 the number of misclassifications using a data-dependent pointer model is 83.1 samples (55.4 %) in average, while the number of incorrectly estimated data using a data-independent pointer model is slightly lower: 82.9 samples or 55.27 % in average. In several cases the number of incorrectly classified samples are the same for both models.

From the results obtained we can state that the difference in estimation using a pointer model with $z(t)$ and without $z(t)$ is really infinitesimal and both models performs the same.

More detailed information can be seen in table 33. The table lists the average values and standard deviation for misclassifications using different noise ratio. It comprises the number of incorrectly classified samples both for a data-dependent pointer model and a pointer model without $z(t)$. Green cells in the table represent the lower number of incorrectly estimated samples using a pointer model with $z(t)$ compared with the number of misclassifications using a pointer model without $z(t)$. The results when a data-dependent model has more wrong classifications are coloured in red. The equal number of incorrectly classified samples for both models are in blue.

Table 33: Classification results for state-space components: average and standard deviation

	rv=0.001		rv=0.1		rv=1		rv=5		rv=10	
	with z(t)	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**	**with z(t)**	**without z(t)**
rw=0.001	0.20±0.63	0.30±0.48	5.20±4.16	5.60±4.45	37.40±3.50	38.10±3.41	65.40±6.08	65.70±5.81	79±5.38	78.70±5.46
rw=0.1	4.50±2.27	4.50±1.78	10.20±2.86	10.80±3.12	43±5.10	43.30±5.06	68.40±5.06	68.80±4.83	77.70±6.73	77.60±6.33
rw=1	32.50±3.63	32.30±3.80	37.40±5.91	38.10±5.99	53.20±4.21	53.20±4.32	69±5.93	68.70±6.13	77.90±3.35	78.20±4.10
rw=5	65.30±5.14	65.40±5.28	64.80±4.26	65±4.45	69.10±5.22	69.40±5.17	76.10±5.82	76.50±5.50	79±6.41	79.10±6.17
rw=10	77.40±5.60	77.50±6.35	75.50±2.99	75.30±3.27	76.20±4.37	76±4.27	79.80±5.45	79.50±5.54	83.10±4.20	82.90±3.67

3.5 Normal components

3.5.1 Data simulation

To perform simulation the parameters for three components and for a pointer model have been defined. The total number of simulated data is set to 500, including data for component model $y(t)$, data for pointer model $c(t)$, discrete values for a pointer model $z(t)$. The simulated data are then used for estimation for both static data-dependent and data-independent pointer model.

During simulation we have change the values of parameters in a way to increase the uncertainty and to make the estimation a challenging task for a proposed algorithm. For each combination of changed parameters ten experiments were performed. The total number of experiments for normal components is 80.

Some examples of simulated data can be seen in the following figures.

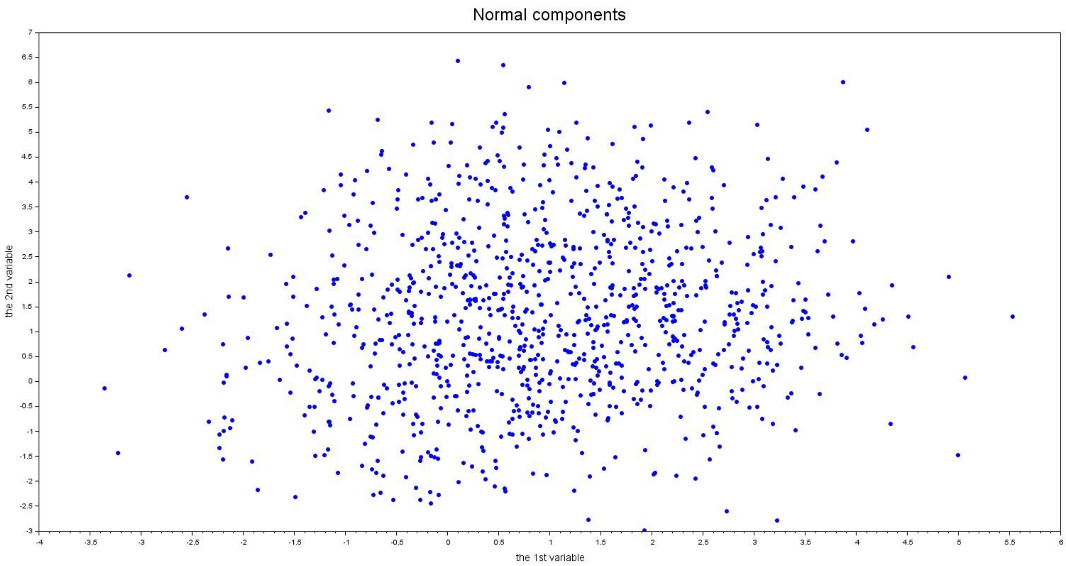


Figure 36: Simulation of state-space components (the pointer model is almost deterministic)

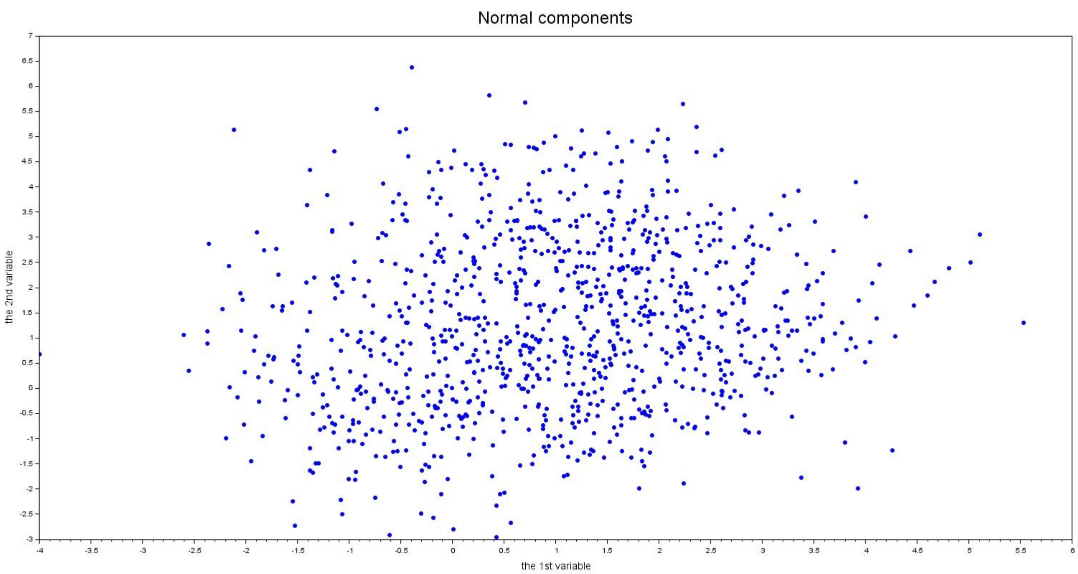


Figure 37: Simulation of state-space components (great amount of noise)

3.5.2 Results

The results of the performed simulation and estimation both for static data-dependent and data-independent pointer models are presented in the tables below. The tables show the number of incorrectly classified samples with a data-dependent pointer (with $z(t)$) via the number of misclassifications with a data-independent pointer (without $z(t)$).

The first table is for the parameters:

Table 34: Parameter values

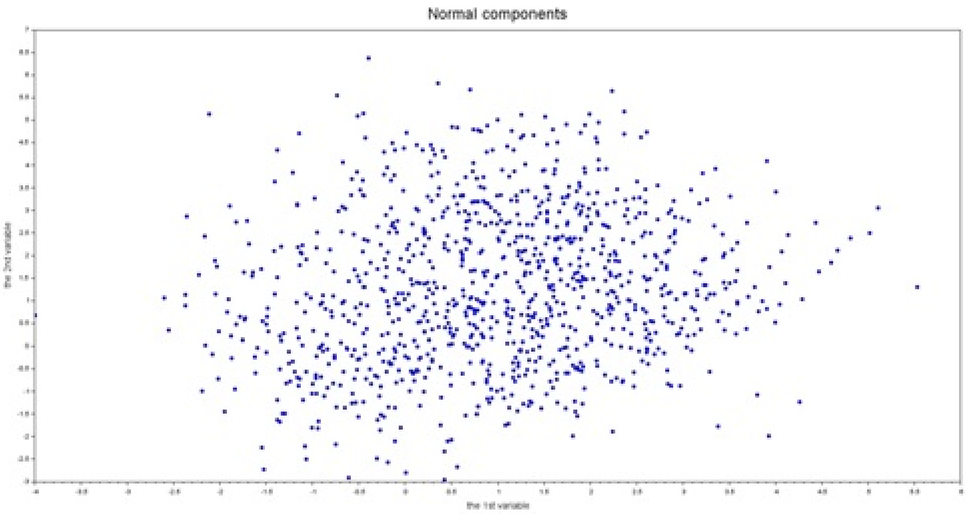
z_t	c_t		
	1	**2**	**3**
1	0.998	0.001	0.001
2	0.001	0.998	0.001
3	0.001	0.001	0.998

Table 35: Misclassified samples

№	Number of misclassified samples			
	with z(t)	**in [%]**	**without z(t)**	**in [%]**
1	3	0.6	325	65
2	2	0.4	324	64.8
3	2	0.4	325	65
4	2	0.4	325	65
5	154	30.8	324	64.8
6	3	0.6	325	65
7	3	0.6	324	64.8
8	3	0.6	324	64.8
9	2	0.4	325	65
10	2	0.4	324	64.8

Examples of simulation and estimation for the above mentioned parameters are shown in the following table:

Table 36: Classification examples

<table>
<tr><th>with z(t)</th><th>without z(t)</th></tr>
<tr><td colspan="2">

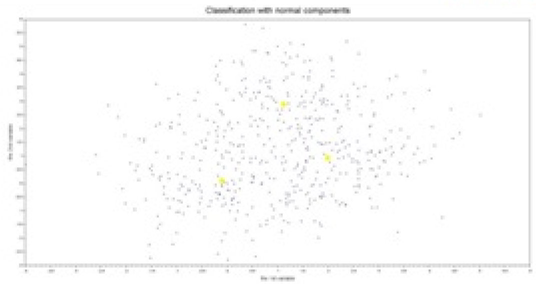
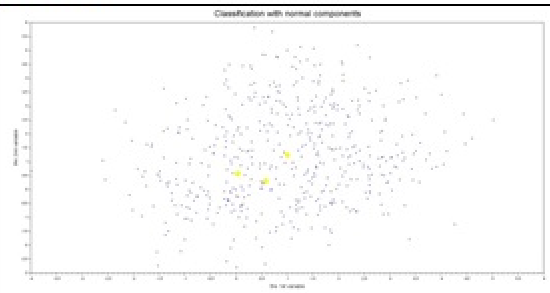
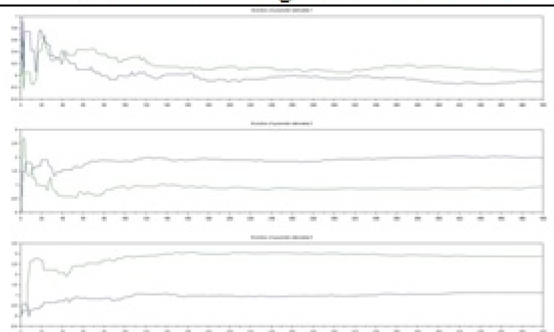
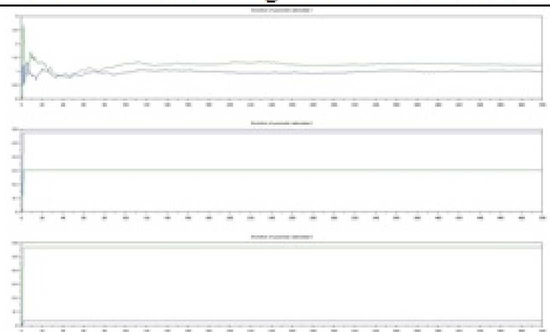
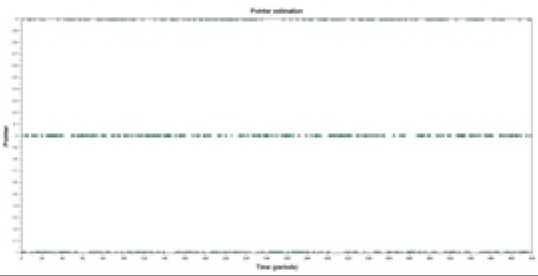
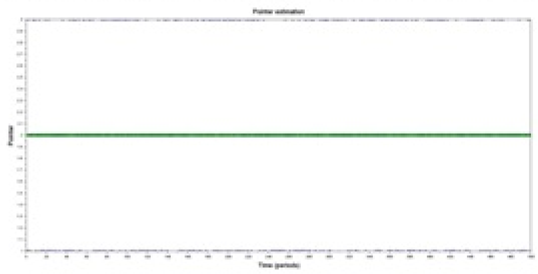
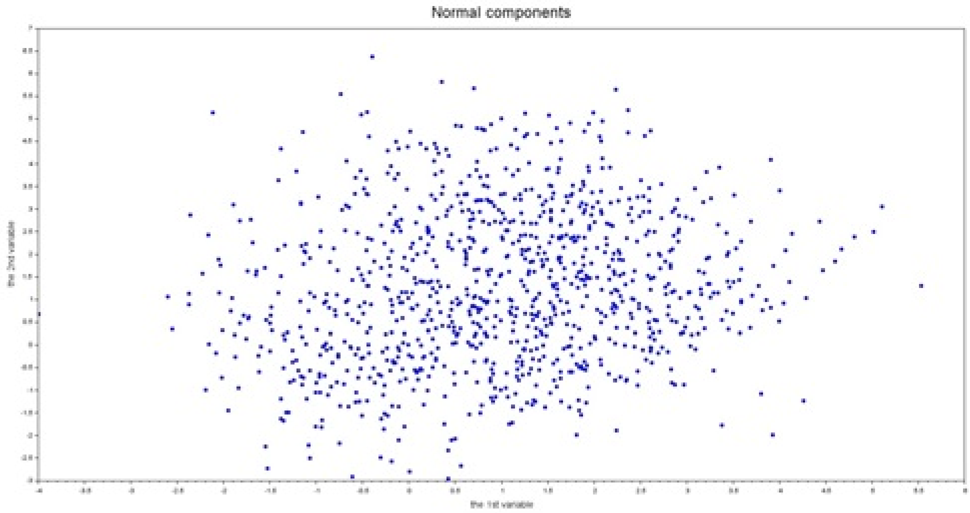
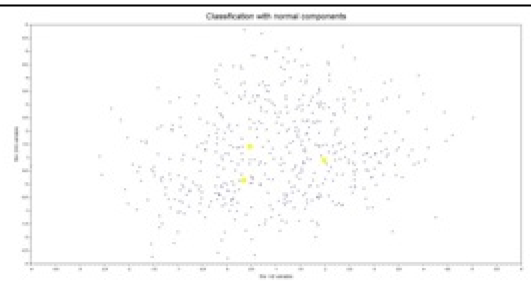
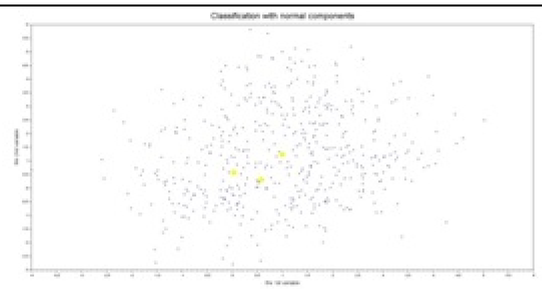
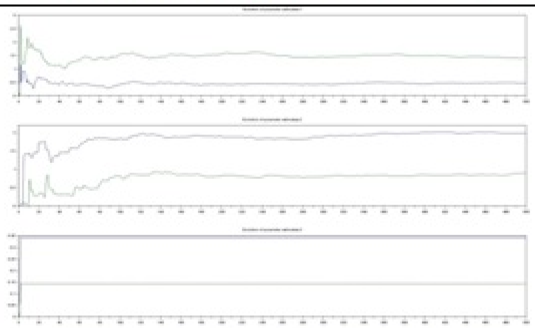
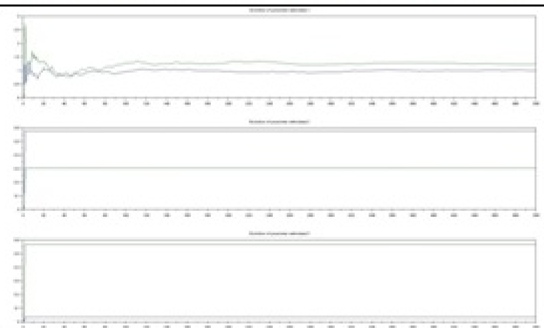
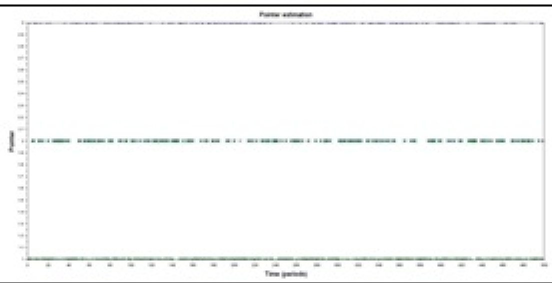
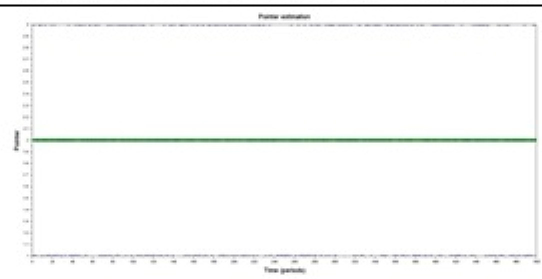
Simulation results</td></tr>
<tr><td>
The number of wrong classifications is 3</td><td>
The number of wrong classifications is 325</td></tr>
<tr><td></td><td></td></tr>
<tr><td></td><td></td></tr>
</table>

Table 37: Classification examples

<table>
<tr><th>with z(t)</th><th>without z(t)</th></tr>
<tr><td colspan="2">
</td></tr>
<tr><td>

The number of wrong classifications is 154</td><td>

The number of wrong classifications is 324</td></tr>
<tr><td></td><td></td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

The next table is for the parameters:

Table 38: Parameter values

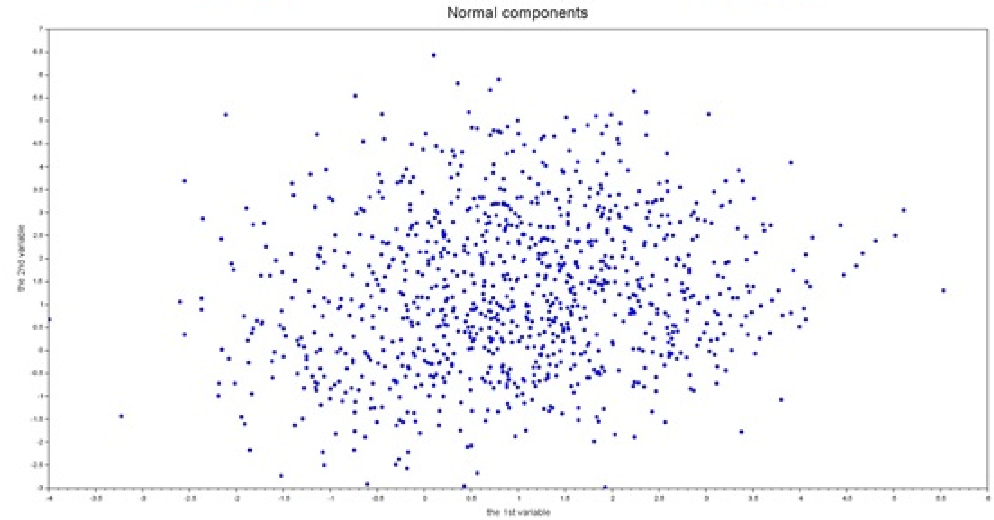
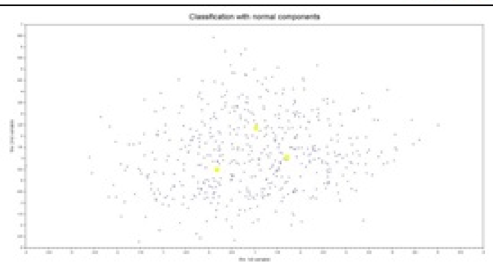
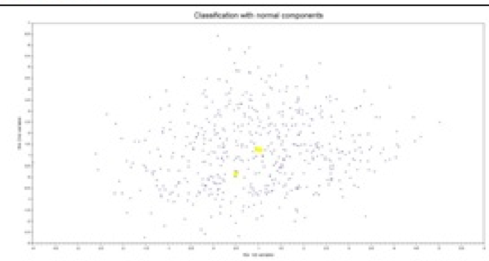
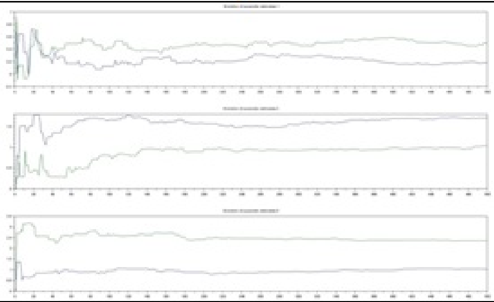
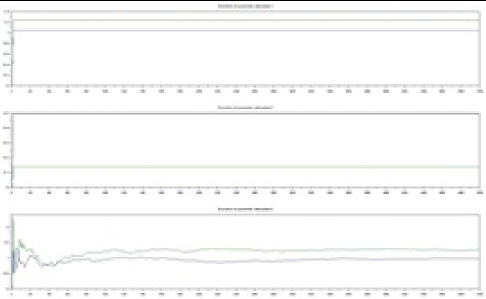
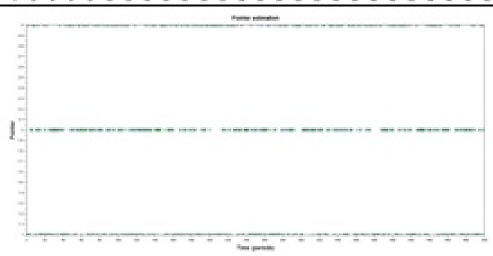
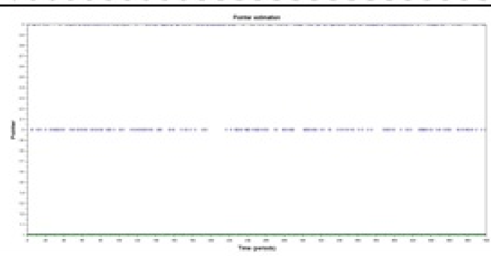
z_t	c_t		
	1	**2**	**3**
1	0.8	0.1	0.1
2	0.1	0.8	0.1
3	0.1	0.1	0.8

Table 39: Classification results

№	Number of misclassified samples			
	with z(t)	**in [%]**	**without z(t)**	**in [%]**
1	100	20	324	64.8
2	101	20.2	324	64.8
3	100	20	324	64.8
4	100	20	324	64.8
5	100	20	324	64.8
6	200	40	324	64.8
7	101	20.2	324	64.8
8	101	20.2	324	64.8
9	101	20.2	324	64.8
10	100	20	324	64.8

Comparison of the results for data-dependent and data-independent pointer models with the above mentioned initial parameters are shown in the following table:

Table 40: Classification examples

with z(t)	without z(t)
	
Simulation results	
 The number of wrong classifications is 100	 The number of wrong classifications is 324
	
	

The third table is for the parameters:

Table 41: Parameter values

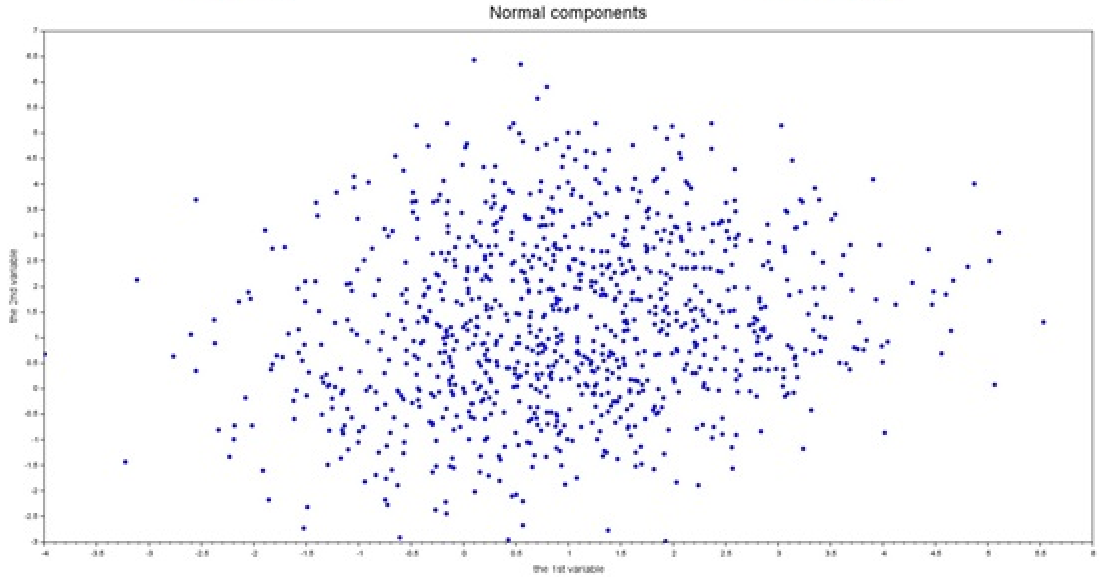
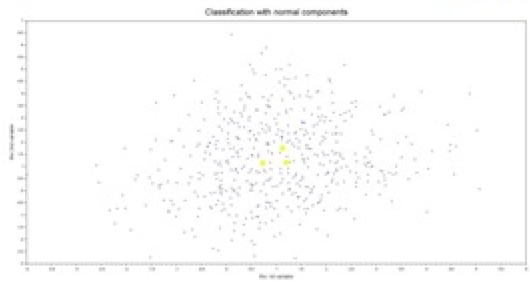
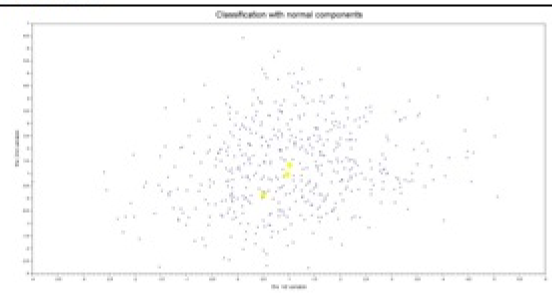
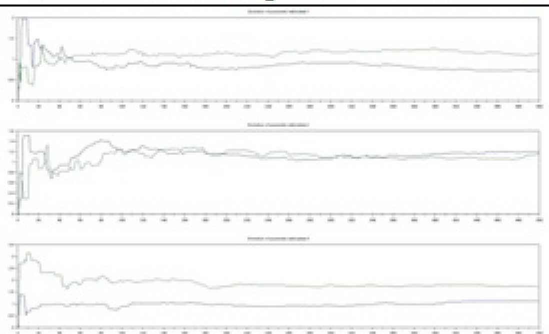
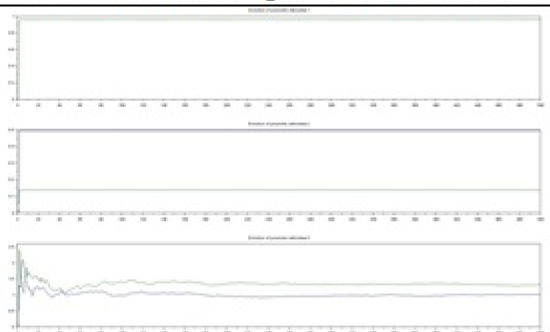
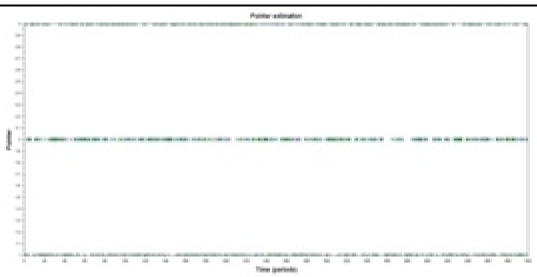
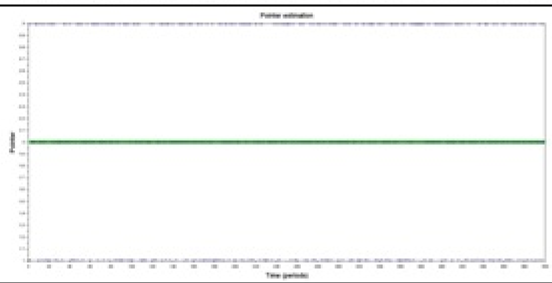
z_t	c_t		
	1	**2**	**3**
1	0.5	0.25	0.25
2	0.25	0.5	0.25
3	0.25	0.25	0.5

Table 42: Classification results

№	Number of misclassified samples			
	with z(t)	**in [%]**	**without z(t)**	**in [%]**
1	258	51.6	329	65.8
2	259	51.8	330	66
3	258	51.6	329	65.8
4	258	51.6	330	66
5	258	51.6	330	66
6	294	58.8	329	65.8
7	259	51.8	330	66
8	259	51.8	329	65.8
9	259	51.8	329	65.8
10	258	51.6	330	66

Some examples of the results for used parameters are listed in the following table:

Table 43: Classification examples

with z(t)	without z(t)
	
Simulation results	
 The number of wrong classifications is 258	 The number of wrong classifications is 329
	
	

The last table is for the parameters:

Table 44: Parameter values

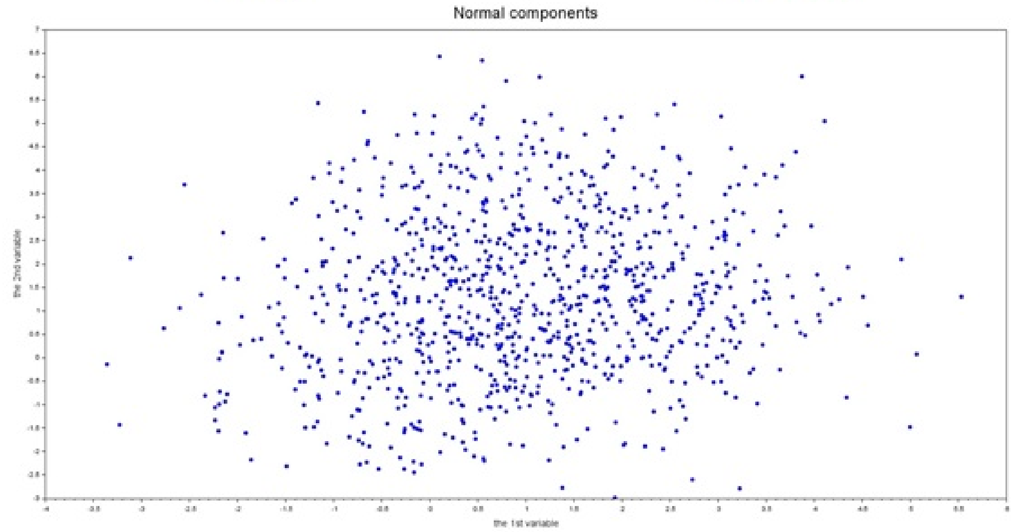
z_t	c_t		
	1	**2**	**3**
1	0.1	0.45	0.45
2	0.45	0.1	0.45
3	0.45	0.45	0.1

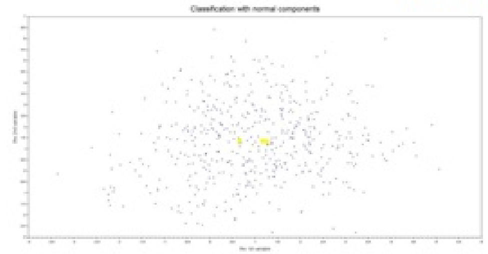
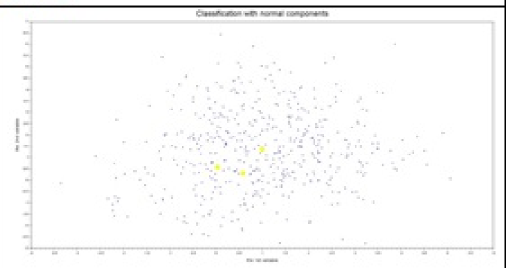
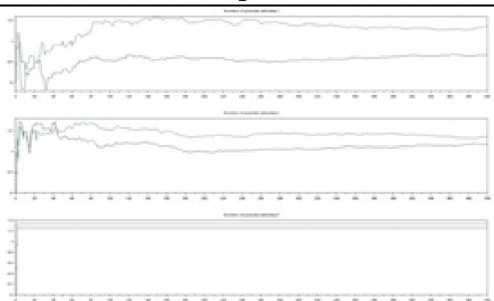
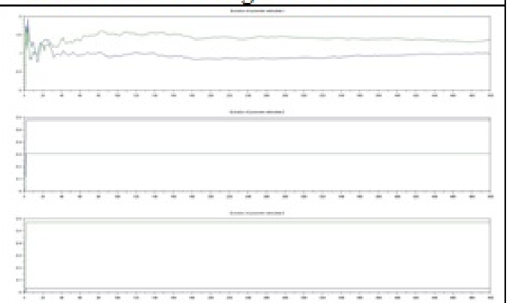
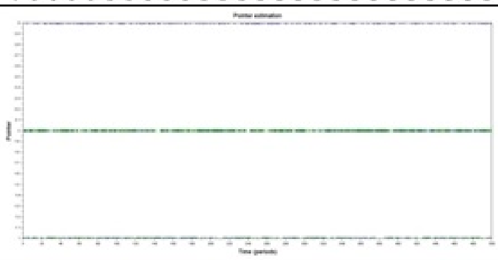
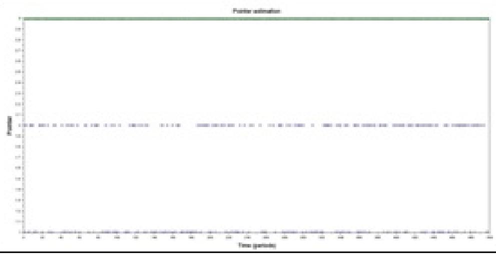
Table 45: Classification results

№	Number of misclassified samples			
	with z(t)	**in [%]**	**without z(t)**	**in [%]**
1	279	55.8	322	64.4
2	279	55.8	322	64.4
3	279	55.8	321	64.2
4	279	55.8	322	64.4
5	279	55.8	322	64.4
6	279	55.8	322	64.4
7	279	55.8	321	64.2
8	279	55.8	322	64.4
9	280	56	322	64.4
10	279	55.8	322	64.4

The following figures present the results of simulation:

Table 46: Classification examples

<table>
<tr><th>with z(t)</th><th>without z(t)</th></tr>
<tr><td colspan="2">

Simulation results</td></tr>
<tr><td>
The number of wrong classifications is 279</td><td>
The number of wrong classifications is 322</td></tr>
<tr><td></td><td></td></tr>
<tr><td></td><td></td></tr>
</table>

The results of estimation and classification are discussed in section 3.5.3.

3.5.3 Discussion

From the tables in previous section it is obvious that the greater uncertainty is the worse results are shown both by data-dependent and data-independent pointer model. It should be noted that the data-dependent pointer model gives much better results in comparison with the data-independent pointer model (without $z(t)$). However, with increasing noise the accuracy of estimation is going down and approaches to that of the data-independent pointer model.

To make the results more obvious table 47 shows the average value and standard deviation for classification results with different values of parameters both for data-dependent and data-independent pointer models. In all cases the data-dependent pointer model shows more accurate estimation. For parameters [0.998 0.001 0.001; 0.001 0.998 0.001; 0.001 0.001 0.998] the number of wrong classifications is generally 2-3 samples (0.4 % to 0.6 %), but one experiment for these values of parameters shows the number of wrong classifications as being 154 or 30.8 %. This is the only one case from 10 experiments. This extreme value is not included in calculation of mean and standard deviation. The number of misclassifications for data-independent pointer model in this case is from 324 to 325 samples (64,8 % to 65 %).

For parameters being [0.8 0.1 0.1; 0.1 0.8 0.1; 0.1 0.1 0.8] the situation is worse. For a data-dependent pointer model the number of incorrectly classified samples is generally from 100 to 101 (from 20 % to 20.2 %). Again one experiment shows the greater number of classifications than others 200 (40 %). This value was not taken into account when calculating the average and standard deviation. The number of misclassifications using a data-independent pointer model remains the same as in previous case: 324 samples (64,8 %).

Setting parameters to values [0.5 0.25 025; 0.25 0.5 0.25; 0.25 0.25 0.5] we obtain 262 wrong classifications in average (52,7 %) using a data-dependent pointer model and 329.5 samples in average (65,8 %) using a data-independent pointer model.

For the last case the number of incorrectly estimated samples is 279,1 (55,82 %) and 321.8 (64,36 %) samples in average for a data-dependent and data-independent model resp.

Table 47: Classification results for normal components: average and standard deviation

Parameters	Misclassifications using model with z(t)	Misclassifications using model without z(t)
[0.998 0.001 0.001; 0.001 0.998 0.001; 0.001 0.001 0.998]	2.4 ± 0.52 (47.93)	324.5 ± 0.53
[0.8 0.1 0.1; 0.1 0.8 0.1; 0.1 0.1 0.8]	100 ± 0.52 (31.49)	324 ± 0
[0.5 0.25 0.25; 0.25 0.5 0.25; 0.25 0.25 0.5]	262 ± 11.26	329.5 ± 0.53
[0.1 0.45 0.45; 0.45 0.1 0.45; 0.45 0.45 0.1]	279.1 ± 0.32	321.8 ± 0.42

4 Conclusion

The presented report is devoted to the analysis of a data-dependent pointer model, whether it brings some advantages in comparison with a data-independent pointer model at simulation and estimation of components referring to different types of distribution, including categorical, uniform, exponential and state-space components for a dynamic data-dependent model, and normal components for a static data-dependent pointer model. For these purposes data from corresponding distributions were simulated in SciLab: 500 samples for uniform, exponential and normal components, and 150 samples for categorical and state-space components. During simulation the parameters of component models and pointer models were defined and changed to add or reduce noise and to change the distance between components and, thus, to analyze, in what situations the data-dependent pointer model performs better. For each type of specified component distributions 200-360 experiments were made both for the data-dependent and data-independent pointer models. Based on the results of wrong classifications/predictions for different variations of the parameters the average of incorrectly estimated samples and standard deviation were calculated. According to the obtained results - both the numbers of misclassified samples and graphs - it was concluded that the present algorithm with a dynamic data-dependent pointer model did not improve the accuracy of estimation. Thus, for categorical and uniform components the algorithm with a data-dependent pointer model has more or less the same performance. The state-space components with a data-dependent pointer model demonstrate slightly better results than that with a data-independent pointer model. The worst results using a data-dependent pointer model are shown for exponential components, where the algorithm with a data-dependent pointer model fails to present good results almost in all cases apart for the cases when the components are really far from each other and their classification is obvious even at first sight. Taking into account all obtained information it can be stated that the present algorithm with a dynamic data-dependent pointer model (with $z(t)$) performs similarly well as the one without $z(t)$.

However, a static data-dependent pointer model using expert knowledge in the beginning of estimation shows much better results than a data-independent pointer model. Though with growing noise the accuracy of estimation is decreasing, the results are still better while using a data-dependent pointer model in this case. It can be explained that in the case of the dynamic data-dependent pointer model the dependence on the previous pointer value is dominating in comparison with the discrete data in the condition. It seems that for such a configuration of the model the discrete data don't bring any new strong information. Unlike this, the static data-dependent pointer can use advantages of this dependence. It would be highly recommended to try this algorithm for all other distributions and to see what the results could be.

Acknowledgements

The research was supported by project GAČR GA15-03564S.

References

[1] Bouveyron, C., Brunet-Saumard, C. (2014). Model-based clustering of high-dimensional data: A review. *Computational Statistics & Data Analysis.* 71(0), p. 52–78.

[2] Zeng, H., Cheung, Y. (2014). Learning a mixture model for clustering with the completed likelihood minimum message length criterion. *Pattern Recognition.* 47(5), p. 2011–2030.

[3] Ng, S.K., McLachlan, G.J. (2014). Mixture models for clustering multilevel growth trajectories. *Computational Statistics & Data Analysis.* 71(0), p. 43–51.

[4] Gupta, M. R. , Chen, Y. (2011). Theory and use of the EM method. In: *Foundations and Trends in Signal Processing*, vol. 4, 3, p. 223–296.

[5] Boldea, O., Magnus, J. R. (2009). Maximum likelihood estimation of the multivariate normal mixture model, *Journal of The American Statistical Association*, vol. 104, 488, p. 1539–1549.

[6] Wang, H.X., Luo, B., Zhang, Q. B., Wei, S. (2004). Estimation for the number of components in a mixture model using stepwise split-and-merge EM algorithm,

[7] McGrory, C. A., Titterington, D. M. (2009). Variational Bayesian analysis for hidden Markov models, *Australian & New Zealand Journal of Statistics*, vol. 51, p. 227–244.

[8] V. Šmídl, A. Quinn, The Variational Bayes Method in Signal Processing, Springer-Verlag Berlin Heidelberg, 2006.

[9] Z. Ghahramani, G. E. Hinton, Variational learning for switching state-space models, Neural Computation, vol. 12, 4 (2000), p. 831–864.

[10] S. Frühwirth-Schnatter, Finite Mixture and Markov Switching Models, Springer-Verlag New York, 2006.

[11] A. Doucet, C. Andrieu, Iterative algorithms for state estimation of jump Markov linear systems, IEEE Transactions on Signal Processing, vol. 49, 6 (2001), p. 1216–1227.

[12] R. Chen, J. S. Liu, Mixture Kalman filters, Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 62 (2000), p. 493–508.

[13] V. Peterka, Bayesian system identification, in: Trends and Progress in System Identification (ed. P. Eykhoff), Oxford, Pergamon Press, 1981, p. 239–304.

[14] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, L. Tesař, Optimized Bayesian Dynamic Advising: Theory and Algorithms, Springer-Verlag London, 2006.

[15] M. Kárný, J. Kadlec, E.L. Sutanto, Quasi-Bayes estimation applied to normal mixture, in: Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing (eds. J. Rojíček, M. Valečková, M. Kárný, K. Warwick), CMP'98 /3./, Prague, CZ, 07.09.1998–09.09.1998, p. 77–82.

[16] I. Nagy, E. Suzdaleva, M. Kárný, T. Mlynářová, Bayesian estimation of dynamic finite mixtures, Int. Journal of Adaptive Control and Signal Processing, vol. 25, 9 (2011), p. 765–787.

[17] Suzdaleva, E., Nagy, I., Mlynářová, T. (2015). Recursive Estimation of Mixtures of Exponential and Normal Distributions. In: *Proceedings of the 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Warsaw, Poland, September 24–26, p.137–142.

[18] Nagy, I., Suzdaleva, E., Mlynářová, T. (2016). Mixture-based clustering non-gaussian data with fixed bounds. In: *Proceedings of the IEEE International conference Intelligent systems IS'16*, Sofia, Bulgaria, September 4–6, accepted.

[19] Nagy, I., Suzdaleva, E., (2013). Mixture estimation with state-space components and Markov model of switching. *Applied Mathematical Modelling*, vol. 37, 24, p. 9970–9984.