International Conference and Workshop on Basic and Applied Sciences proceeding
October $16^{th} - 17^{th}$ 2015, Surabaya Indonesia.

THE NUMERICAL SIMULATIONS OF INVERSE PROBLEMS ON THE PARAMETER ESTIMATION

Julan HERNADI
Department of Mathematics Education, Muhammadiyah University of Ponorogo
julan_hernadi@yahoo.com

Abstract. The problem of determination an estimator through sample data is a part of the inverse problems. Generally, the inverse problem has no solution in the ordinary sense since the known data have been contaminated by noises. The minimizer of the least square functional is usually taken as the solution of the invers problem.

The optimal design on parameter estimation uses the Fisher information matrix (FIM) as a tool for optimal criteria that minimizes some cost functional over set of FIM's. The consideration is based on the Cramer-Rao lower bound inequality that can only be attained by the invers of FIM.

This article demonstrates how to implement the inverse problems in connection with parameter estimation numerically where the set of noises is generated independenty during the experiment. The numerical simulation is applied to a distributed parameter system of parabolic equation to find the optimal sensor locations for the parameter. The simulation is also carried out to a model of dynamical system to obtain the optimal time for measurements. The accuracy of estimators are compared to the prior supposed nominal paramaters and the variance of estimators are contrasted with the lower bound of Cramer-Rao inequality through the functional value acting on FIM's.

Keywords: Optimal design, FIM, parameter estimation, inverse problem.

1. INTRODUCTION

In general, two approaches are often used to study the real world fenomena, namely through their mathematical representation and by making replication of their behavior. On the mathematical representation, a relevant mathematical model is built dan studied. In most cases, the mathematical model involves some parameters that require to be estimated through observation data. This stage is known as the calibration model.

At the beginning, the problem of parameter estimation appeared in statistics where the functional relationship among variables is presumed to have an explicit representation like linear, quadratic, exponential, and some other, and the set of parameters in the statistical model is assumed arising from a population with certain distribution probability.(Goodwin [7]).

In latest development, the explicit relationships were not adequate anymore to deal problems in science and engineering. The implicit relationship among variables like differential equations subjected to initial or boundary value conditions, integral

2010 Mathematics Subject Classification: 37M05, 62B15, 62G08, 65C60, 65K10, 93B51, 94A20.

equations, or integro-differential equations had attracted the attention many mathematicians and applied scientists (Vogel, [13], Doicu at. al. [6], Muller [10]; Aster at. al. [1]).

The abstract formulation of parameter estimation (identification) can be described by the following operator equation.

$$G(\theta) = d, \tag{1.1}$$

where G denotes an operator, θ constitutes the parameter vector characterizing the model and d represents the observation data that might be a function of time or a discrete set of observations. Two problems in the parameter identification are the forward problem and the inverse problem. On the forward problem, the paramemeter θ is given, the output model is determined. In this problem, we are faced with problem of solving an operator equation. The inverse problem is adressed to obtain an estimator for parameter θ through the observation data d. The inverse problem is generaly ill-posed because of the noise disturbance most probably accomadated in data observation due to reading mistaken of instrument or rounding error of numerical data. This is the main obstacle in solving the inverse problem in connection with parameter estimation. This situation is represented as

$$\begin{aligned} \text{data} &= \text{clean observation+noise component} \\ d &= G(\theta_{true}) + \epsilon \qquad (1.2) \\ &= d_{true} + \epsilon \qquad (1.3) \end{aligned}$$

where $d_{true} = G(\theta_{true})$, i.e. fulfils exactly equation (1.1) whenever $\theta = \theta_{true}$. This exact parameter is sometimes called the nominal or natural parameter and its exsitence is assumed naturally. Usually, the nominal parameter is not available explicitly, instead, it must be estimated by an estimator obtained through the observation data (sample).

An intuitive consideration in choosing the optimal sample that a sample containing much information for some parameter is believed will result a better estimator in the sense more accurate and more reliable. Accurate refers the closeness to the nominal parameter, while reliability corresponds to the small varians. The premis used on paramater estimation is that the information content on the parameter θ may vary considerably from one time measurement to another (Banks, at. al. [4])

One of the parameter model that frequently appears on the applied sciences is the system of differential equations which models some physcial, sociological or biological phenomenon(Bank at. al. [4, 5]).

$$\dot{x}(t) = g\left(t, x(t), \theta\right), t \geq t_0, x(t_0) = x_0(\theta), \tag{1.4}$$

where $x(t) \in \mathbb{R}^n$ denotes the vector of state variables and $\theta \in \mathbb{R}^p$. This model is a kind of the lumped-parameter system (LPS). For a given the admisible parameter θ, the solution $x = x(t, \theta)$ to the initial value problem (1.4) can be obtained. Sometimes the model output emerges in the form of functional with respect the state variable $x(t, \theta)$, i.e. $f(t, \theta) := h(t, x(t, \theta), \theta)$. Consequently, the model output (1.4) can be represented as

$$z(t) = f(t, \theta), t \geq t_0. \tag{1.5}$$

The distributed-parametric system (DPS) which is a model that depends on both the time variable and the spatial variables also involves some parameters that require to be estimated. One of the DPS model is represented by some partial differential equation subjected to some intial and boundary values as given in (Ucinski [12]):

$$\frac{\partial y}{\partial t} = \mathcal{F}\left(x, t, y, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial^2 y}{\partial x_1^2}, \frac{\partial^2 y}{\partial x_2^2}, \theta\right), x \in \Omega \subset \mathbb{R}^2, t \in (0, T) \tag{1.6}$$

with boundary conditions $\mathcal{B}\left(x, t, y, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \theta\right) = 0$, $x \in \partial\Omega, t \in (0, T)$ and initial value $y(x, 0) = y_0(x), x \in \Omega$.

In the implementation, this system is solved generally by a semi-dicrete method in which the one variable is discretized, while keeping the another continue. Thus, the model output of (1.5) for LPS in some occasion can be adapted to DPS. By assumption the model (1.4) is representative enough of the real problem, the nominal parameter $\theta_0 \in \mathcal{A}$ is guaranted available, and thus the system output $y(t) = f(t, \theta_0)$ can be regarded as the true output. It is assumed that the measurement $z(t)$ at time t contains the error, beside the true data. It means the measurement consist of two components, the system output (clean data) and the error component (noise).

$$z(t) = f\left(t, \theta_0\right) + \epsilon(t),\ t \in [0, T]. \tag{1.7}$$

The random proccess of measurement error $\epsilon(t)$ is assumed to have zero mean, time-dependent varians $\sigma^2(t)$, independent within the measurements.

In this paper, the inverse problem is simulated to estimate the nominal parameter θ_0 through observation data. To asses the sample quality, the estimator variability based on that sample is calculated from a series of experiments where the noises are generated independetly during the experiment. Finally, the estimator $\hat{\theta}$ for θ_0 is taken as the mean of estimators which were obtained from each eksperiment as done in the Monte Carlo simulation.

2. THE OPTIMAL DESIGN METHOD

Considering the output model $f(t, \theta)$ and the statistical model $z(t)$ (observation proccess), the objective function is given by a generalized functional error as

$$J(y, \theta) := \int_0^T \frac{1}{\sigma^2(t)} \left(y(t) - f(t, \theta)\right)^2 \, dP(t) \tag{2.1}$$

where P denotes a general measure defined on σ-algebra of $[0, T]$. In particular P is the probability measure. Let $\tau = \{t_1, \cdots, t_N \,:\, t_1 < t_2 < \cdots < t_N\}$. The following Dirac measure δ_{t_i} is one of simple but important probability measure.

$$\delta_{t_i}(A) = \begin{cases} 1 & \text{if } t_i \in A \\ 0 & \text{if } t_i \notin A \end{cases}$$

For a given set of points $\tau = \{t_1, t_2, \cdots, t_N\}$, the measure P_τ is defined as

$$P_\tau = \sum_{i=1}^{N} \delta_{t_i}. \tag{2.2}$$

By means of this measure, the cost functional in (2.1) reduces to the discrete version of the weighted least squared functional.

$$J_d(y, \theta) = \sum_{i=1}^{N} \frac{1}{\sigma^2(t_i)} \left(z(t_i) - f(t_i, \theta)\right)^2. \tag{2.3}$$

Although the continuous version (2.1) is possible to be applied on numerical implementation, e.g. by quadrature formula for the integral approximation, but the discrete version is more advantage.

According to Banks at. al [4], the abstract problem of the optimal design is how to choose "the best" measure $\hat{P}$ over $\mathcal{P}[0, T]$ the set of all probability measures on $[0, T]$ in the sense

$$\mathcal{J}\left(F(\hat{P}, \theta_0)\right) = \min_{P \in \mathcal{P}(0,T)} \mathcal{J}\left(F(P, \theta_0)\right), \tag{2.4}$$

where $F(P, \theta_0)$ denotes the Fisher information matrix (FIM) given by

$$F(P, \theta_0) = \int_0^T \frac{1}{\sigma^2(s)} \nabla_\theta^T f(s, \theta_0) \nabla_\theta f(s, \theta_0) dP(s). \tag{2.5}$$

In practice, the probality measure P is characterized by mesh, or points in τ as mentioned before such that the optimal design method is now switched to the optimization problem of choosing a set of poitns $\tau^* = \{t_i^*\}$, $i = 1, \cdots, N$ such that the following condition holds.

$$\mathcal{J}\left(F(\tau^*, \theta_0)\right) = \min_{\tau \in \mathcal{T}} \mathcal{J}\left(F(\tau, \theta_0)\right) \tag{2.6}$$

where $\mathcal{T}$ stands for a set of all meshes $0 \leq t_1 \leq \cdots \leq t_N \leq T$

The discrete version of FIM had been introduced by Thomaseth and Cobelli [11] as

$$F(\tau, \theta_0) = \sum_{i=1}^{N} \frac{1}{\sigma^2(t_i)} \nabla_\theta^T f(t_i, \theta_0)\ \nabla_\theta f(t_i, \theta_0) \tag{2.7}$$

where $\nabla_\theta f := \frac{\partial f(t,\theta)}{\partial \theta}$ denotes the vector gradient of parameters. Throughout this article, the simulations use the discrete version of FIM's as the arguments of the objective functional (2.6). Furthermore, the optimal criteria is based on such optimization problem. Many optimal criteria found in the literature (Ucinski [12], Banks at. al [5]), following are three famous:

- D-optimal, which maximize the determinat of FIM, i.e.

  $$\mathcal{J}(F) := -\ln \det(F)$$

  where det denotes the determinant of matrix.
- E-optimal, which maximize the spectral radius of FIM, i.e.

  $$\mathcal{J}(F) := \lambda_{\max}(F^{-1})$$

  where $\lambda_{\max}$ is the largest eigen value.
- A-optimal or SE-optimal, which maximize the trace of FIM, i.e.

  $$\mathcal{J}(F) := \operatorname{tr}(F^{-1})$$

  where "tr" stands for the trace of matrix, the sum of entries on the main diagonal.

Hence, we have two kinds of optimization problems that must be solved in the parameter identification. First, the optimal sample is determined by taking the minimizer of the functional (2.6) based on the optimal criteria. Second, the estimator of paramater θ is taken as the minimizer of the least squared functional in (2.3). Certainly, The larger sample size and the more parameters involved the more difficult to solve problem numerically due to high dimension. On the other hand, the smaller the sample size gives less information about the parameters obtained through it.

Taking the optimal sampel is similar to minimize the number of the measurements and to fix the best time when the measurements are carried out. It is in line with the choosing the best position for measurement in the case of spatial variable. Banks at. al. [4] used the generalized sensitivity function (GSF) to localize the measurement time at point where magnitude of GSF is closed to 1. The closer GSF to 1, the less information contained in data. The number of parameters can be reduced by keeping the most significant parameters and regarding the others as constants. A method for parameter selection to qualify those parameters was introduced by Banks at. al [2]. The problem to optimize the measurements numbers (sample size) as well as its numerical realization are still open.

This paper represents some algorithms and its numerical realization for simulating the inverse problems on the contex of parameter estimation. The cost functional that measures the discrepancy between model output and observation is minimized. Two examples for implementation are the equation of temperature distribution of a rod thin and the logistic model of Verhulst-Pearl equation.

3. NUMERICAL SIMULATION

The first simulation is a simple case in DPS, the equation of temperature distribution of a rod thin taken from (Ucinski [12]):

Example 3.1. *The temperature distribution $y = y(x,t)$ on a rod thin with assumtion no heat source and both ends with perfect insulated.*

$$\begin{aligned} \frac{\partial y}{\partial t}(x,t) &= \theta \frac{\partial^2 y}{\partial x^2}(x,t), \ x \in (0,1), \ t \in (0,t_f) && (3.1)\\ y(x,0) &= \sin(\pi x) && (3.2) \end{aligned}$$

where the parameter θ stands for the diffusivity of the material forming the rod. Suppose the two ends associate with $x = 0$ and $x = 1$, respectivley, and at time $t = 0$ suddenly placed in contact with ice packs at 0^o and that the temperature at the ends is maintained at all later times. The situation is represented by the boundary conditions

$$y(0,t) = y(1,t) = 0, \ t \in (0,t_f). \tag{3.3}$$

It is deliberate that model was taken trivially so that its analytic solution is available and given by

$$y(x,t) = e^{-\theta \pi^2 t} \sin(\pi x). \tag{3.4}$$

We want to find x^1 the location which contains the most information about parameter θ. Theoretically, it is required to solve the optimization problem with the objective function $\mathcal{J}$ defined on the set of FIM's. At this moment, we do not apply this procedure. Instead, the potential sensor locations are prior defined intuitively to be candidates the optimal sample. In order to make the numerical simulation does work, the spatial variable is discretized and the time variable is kept continuous so that the observation process can be represented as

$$z(t) = y_m(t) + \varepsilon_m(t), \ t \in T := (0,t_f)$$

where

$$\begin{aligned} y_m(t) &= \text{col}\left(y(x^1,t), y(x^2,t), \cdots, y(x^n,t)\right)\\ \varepsilon_m(t) &= \text{col}\left(\varepsilon(x^1,t), \varepsilon(x^2,t), \cdots, \varepsilon(x^n,t)\right). \end{aligned}$$

In this case, $z(t)$ is the observation vector in a space of n-dimenstion, $x^j \in [0,1], j = 1, \cdots, n$ are the sensor locations, $y(x^j,t)$ is assumed as the true output at position x^j dan time t, and $\varepsilon = \varepsilon(x,t)$ is the noisy random process with assumption $E(\varepsilon_m(t)) = 0$, $\text{Var}(\varepsilon_m(t)) = \sigma^2(t)$, $\text{Cov}(\varepsilon_m(t)\varepsilon_m(s)) = \sigma(t)\sigma(s)\delta(t-s)$ for $t \in T$, where δ is the Dirac distribution concentrated at origin.

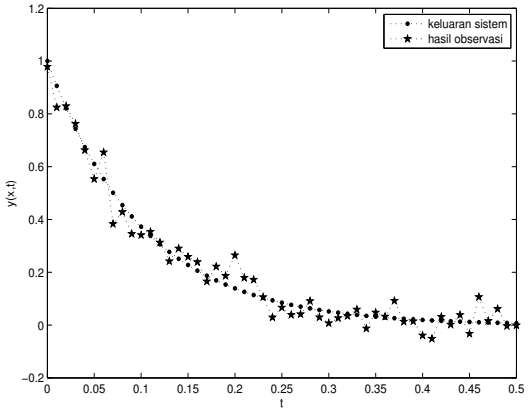
Prior to solving of inverse problem by a certain optimization algorithm, the observation data is irritated by some artificial noise. For each point of optimal sensor x^j, the experiments are carried out repeatedly with different noises. The following algorithm is composed to implement the numerical simulation on computer.

Algorithm 1.

(1) Disceretize the time interval $[0,t_f]$ with sampling or by a uniformly partition with step Δ_t. Let n_t be the lenght or size of time sample.
(2) Define the set of potential sensor location $\Omega_0 = \{x_1, x_2, \cdots, x_N\}$, for instance, taking the equidistance points (uniformly).
(3) For each $x_i \in \Omega_0$:
   (a) Generate the vector of true or system output $y(t, x_i; \theta)$ with respect to $\theta = \theta_{true}$, for t in the time set of sampling, i.e.

$$y_i = (y_i^k | k = 1, \cdots, n_t)$$

   where $y_i^k = y(x_i, t_k)$.



GAMBAR 1. Model output with noise

(b) Generate the noise $\epsilon_i \sim N(0, \sigma^2)$

$$\epsilon_i = (\epsilon_i^k | k = 1, \cdots, n_t).$$

(c) Define the observation data $z_i = (z_i^k | k = 1, \cdots, n_t)$ by taking

$$z_i(t_k) = y(x_i, t_k) + \epsilon_i^k.$$

(d) Define the functional error as the objective function $\mathcal{J}_i : \Theta_{ad} \to \mathbb{R}$ with

$$\mathcal{J}_i(\theta) := \sum_{k=1}^{n_t} |z_i(t_k) - y(x_i, t_k; \theta)|^2.$$

(e) Corresponding to $x_i \in \Omega$, the estimator $\hat{\theta}_i$ is obtained, i.e.

$$\hat{\theta}_i = \text{argmin} \left\{ \mathcal{J}(\theta) | \theta \in \Omega_{ad} \right\}.$$

(f) Do experiment (a)-(e) as L times. The noises are distinguished from one experiment to another.

(g) By applying the Monte Carlo trial, the estimator corresponding to $x_i \in \Omega$ is the mean of L estimators within the experiments.

(h) Calculate the standard error of the estimator to look it's variability.

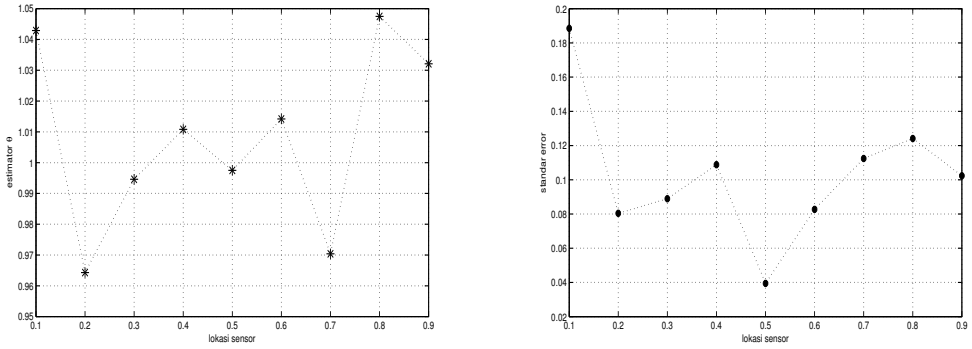
The sensor location x^j corresponding with smallest value of the standard errors qualifies as the most informative sample.

Numerical Experiment 1. For this simulation, we assume $\theta_{true} = 1$, $t_f = 2$, $\epsilon \sim N(0, 0.05)$ is random noises generated by randn on MATLAB. One of the system output at $x = 0.5$ (smooth curve) and the observation data (irregular curve) is represented on Figure 1. It can be seen that all values of the obsevation data are disturbed by noises. The sampling time takes unifomly with $\Delta t = 0.01$. Suppose the optimal sensor locations being observed are given in the set

$$\Omega_0 = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}.$$

The experiment is carried out as $L = 20$ times where the noise is generated independently. The optimization algorithm used the MATLAB optimization toolbox, in particular function fminsearch and fminbnd. Both MATLAB functions produce the same results. In this numerical experiment also tried fminbnd for dealing 1-d optimization problem. The standard error is calculated by the following formula.

$$SE_\theta(x) = \sum_{k=1}^{L} (\hat{\theta}_k(x) - \theta_{true})^2.$$



GAMBAR 2. Estimators at sensor locations and their corresponding standard errors

Figure 2 (left panel) shows the estimator obtained through each sensor location. It is clear that the estimator in connection with $x = 0.5$ is the most accurate. Notice that, $\theta_{true} = 1$ was set as the nominal parameter.

The standard errors for each estimator are shown on Figure 2 (right panel). It is similar to phenomenon in the accuration, the smallest variance is also happened at same point, $x = 0.5$. This result indicates that the location $x = 0.5$ is most optimal, i.e. it contains most information about parameter θ.

Furthermore, we confirm this finding to the theorical background stating that the optimal criteria must be satisfied by this point. By using the elementry calculus, the following FIM is obtained.

$$M(x) = \underbrace{\frac{\pi^4}{\sigma^2} \int_0^{t_f} t^2 e^{-2\theta\pi^2 t}\, dt}_{>0} \sin^2(\pi x).$$

Because of only single point considered, we can treat the FIM as a scalar so that the objective function is trivial. Thus, we can take x^j as the maximum value of FIM instead of using the optimal criteria such E-optimal or A-optimal. This matrix reaches maximum at $\pi x = \pi/2$ or $x = 0.5$. Perfectly, the numerical simulation result and the background theoretical are agreed.

For the second simulation, we use the model of dinamical system of logistic population growth.

Example 3.2. *Model of Verhulst-Pearl logistic describes the dynamical growth of one population which has the intrinsic growth and growth with saturation due to the carrying capacity.*

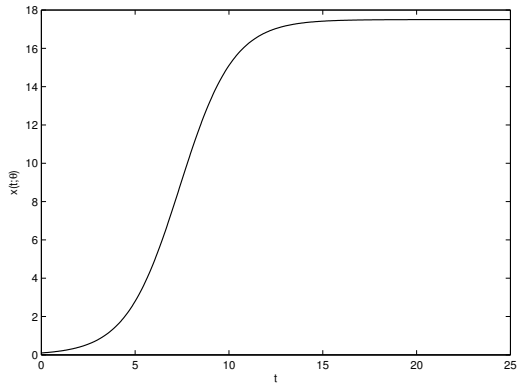
$$\dot{x}(t) = r\, x(t) \left(1 - \frac{x(t)}{K}\right), x(0) = x_0 \tag{3.5}$$

where K denotes the carrying capacity, r stands for the intrinsic growth, and x_0 indicates the initial size of population.

The corresponding parameters system given by $\theta := (K, r, x_0)$. For simplicity, the simulation considers only two parameters, namely r and x_0. In fact, the state variable $x(t, \theta)$ was less sensitif with respect to parameter K than the others [9]. For that reason, K is taken as constant $K = 17.5$. The analitic solution of the model is given explicitly by

$$x(t) = \frac{17.5}{1 + (\frac{17.5}{x_0} - 1)\exp(-rt)}. \tag{3.6}$$

The profile of solution for $x_0 = 0.7$ dan $r = 0.1$ is shown on Figure 3.



GAMBAR 3. Graphic of solution to logistic model for $x_0 = 0.7$ and $r = 0.1$

A relevant problem in connection with the parameter estimation is how to choose the measurements time so that the collected data contains as much as possible the information about parameters x_0 and r. Specifically,

- Where is the interval (duration) for measurements,
- How many measurements must be taken, and
- How is the distribution of the measurements.

The inverse problems is solved for each given a set of measurements data (observation process) repeatedly. On each repetition, a set of noises is disturbed to measurement data. The following algorithm required to realize the numerical simulation.

Algorithm 2.

(1) Define the collection of measurements set or a potential samples $\Omega_0 := \{\tau_1, \tau_2, \cdots, \tau_N\}$ where $\tau_i = \{t_{i1}, t_{i2}, \cdots, t_{in}\}$ is the set of measurement times. In this case, each sample consists of the same number of points, i.e. n.
(2) Fix L, the number of repetitions.
(3) Generate M, the randomly matrix with size $L \times n$ where its rows are the vectors of noises $\varepsilon \sim N(0, \sigma^2)$.
(4) For each $\tau_i \in \Omega_0$:
   (a) Define the true output $x(t;\theta)$ with rescpect to $\theta = \theta_{true}$ so that $x = (x(t;\theta_{true})|t \in \tau_i)$.
   (b) For each $k = 1, \cdots, L$, define the observation data $z_i^k(t)$ composed of the true output component and the noise component $z_i^k(t) = x(t;\theta_{true}) + \epsilon^k$ where $t \in \tau_i$ and ε^k denotes the random vector of k^{th} row of matrix M.
      (i) Define the functional $\mathcal{J}_i^k : \Theta_{ad} \to \mathbb{R}$ with

$$\mathcal{J}_i^k(\theta) := \sum_{t \in \tau_i} |z_i^k(t) - x(t;\theta)|^2.$$

The following minimizer is obtained.

$$\hat{\theta}_i^k = \text{argmin} \left\{ \mathcal{J}_i^k(\theta) | \theta \in \Omega_{ad} \right\}.$$

      (ii) Finally, the estimator corresponding to the measurement time of $\tau_i = \{t_{i1}, t_{i2}, \cdots, t_{in}\}$ is defined as

$$\hat{\theta}_i := \frac{1}{L} \sum_{k=1}^{L} \hat{\theta}_i^k.$$

(c) Calculate the standard error of this estimator to know its variability by formula

$$SE_\theta(\tau_i) = \sum_{k=1}^{L} (\hat{\theta}_k(\tau_i) - \theta_{true})^2.$$

where $\hat{\theta}_k(\tau_i)$ denotes the estimator obtained through sample τ_i at kth repetition.

Totally, there are N vector of estimators corresponding to N measurements sample. The sample correspoding to the estimator with minimum standard error dan most accurate is appointed as the best sample.

Numerical experiment 2. In the implementation, the experiment strategic yields a constrained optimization as introduced in Hernadi [8] and Banks at. al. [5]. A slightly different from precursors, this time a 13 intervals of potential measurements are given in advance as shown in Tabel 1. On each interval, 5 points are taken uniformly where both ends interval are included. The set of measurements (samples) which derive from each interval is used to estimate θ.

TABLE 1. Some intervals for measurement

No	Interval	No	Interval	No	Interval	No	Interval	No	Interval
1	$[0, 4]$	4	$[12, 16]$	7	$[8, 16]$	10	$[10, 20]$	13	$[12, 20]$
2	$[4, 8]$	5	$[16, 20]$	8	$[16, 22]$	11	$[0, 25]$		
3	$[8, 12]$	6	$[2, 8]$	9	$[0, 10]$	12	$[5, 10]$		

The nominal parameter had been choosen $r = 0.7$ and $x_0 = 0.1$ $L = 50$ and $\sigma = 0.125$ for the standard of deviation. The MATLAB optimization toolbox fminsearch had been applied with equal initial values for all experiments. The result of numerical experiments is sumarized in Tabel 2.

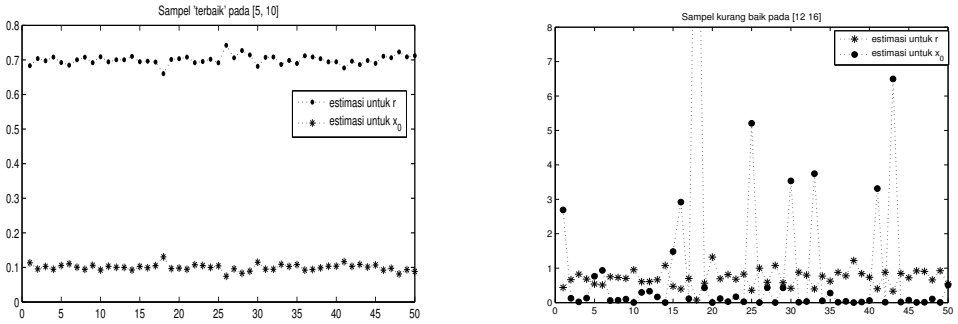
TABLE 2. The estimation and standard errors for each sample

Sampel	Est parameter		Standard error	
	r	x_0	r	x_0
1	0.6954	0.1048	0.4971	0.0495
2	0.6995	0.1005	0.0119	0.0047
3	0.6996	0.1022	0.0278	0.0238
4	0.7075	1.0379	2.7343	390.79
5	1.3641	3.7687	83.300	2696.8
6	0.6985	0.1012	0.1220	0.0049
7	0.7001	0.1018	0.0290	0.0235
8	1.4344	3.6315	84.200	2474.0
9	0.6993	0.1010	0.0112	0.0054
10	0.7068	0.1344	0.3802	0.8510
11	0.6988	0.1039	0.0832	0.0320
12	0.7002	0.1000	0.0084	0.0043
13	0.7378	1.2389	18.5817	18.5910

According to table, it is clear that the worst estimation are given by sample 5 and sample 8, not only a lack of accuracy but also their standar errors are very big. For sample 4 and sample 13, the accuracy is good for parameter r but poor for parameter x_0, while the standard errors are still rather big. The best result is given by sample 12, then follow by sampel 2.

TABLE 3. The values of optimal criteria for each sample

No of Sample	D-optimal	E-optimal	SE-optimal
1	-5.8509	0.8775	0.8808
2	-12.8318	0.0220	0.0221
3	-11.4845	0.0618	0.0619
4	-2.5850	10.4363	10.4435
5	8.4335	3858.4	3859.6
6	-12.3877	0.0248	0.0250
7	-10.6975	0.0865	0.0868
8	8.8811	5131.5	5132.9
9	-9.6012	0.4546	0.4548
10	-6.4705	1.4182	1.4193
11	-9.4171	0.1657	0.1662
12	-13.2531	0.0152	0.0153
13	-1.7538	18.5817	18.5910



GAMBAR 4. Estimator derived from 50 experiments

For a confirmation the result to the theoretical background, the Fisher information matrix (FIM) is calculated for each sample by formula (2.7) with $\sigma = 1$, then the corresponding functional with optimal criteria is evaluated. The computation results are presented in Table 3. Based on thi evaluation, the optimal criteria value of sample 5 and 8 much bigger than the others. Theoretically, the smaller value of optimal criteria the better of estimator as its variance closer to the lower bound of Cramer-Rao inequality. The best two are sample 12 and sample 2 which give small values of the optimal criteria. For instance, the sample 8 is regarded as bad sample as the value 8.4335 for D-optimal and 3.8584×10^3 for E-optimal, meanwhile the sample 12 is considered as the best since the value for D-optimal is -13.2531 and 0.0153 for E-optimal. The results of SE-optimal resemble E-optimal.

The influence of those 50 experiments with respect to the parameter estimation is visualized by two graphs on Figure 4 that compares between the best sample and another. On left panel, the measurements were taken on $[5, 10]$, i.e. sample 12. It can be seen that the estimators for both paramaters are very settle. This behaviour indicates the sampel contains much information about the paramemeter. However, on the right panel the measurements were taken on the interval $[12, 16]$, i.e. sample 4. Although the accuracy is good enough as 0.7075 for r but it come worst for x_0 as 1.0379. Also, their variability are very big, in particular for x_0. This sample is regarded as the less informative for parameter being estimated than previous one.

4. CONCLUDING REMARKS

The numerical simulation of inverse problem has been applied to the problem of finding the optimal location for sensor and the optimal time for measurement. According to simulations, the optimal criteria does play important role to obtain the optimal sample in the sense it contains much information about the parameters being estimated. Although the lower bound of Cramer-Rao inequality is almost impossible to be achieved, it is suggested to have a sample with small optimal criteria value, the smaller this value the better sample quality for estimating in the sense more accurate and more resistence due to small disturbance on the measurement data.

This simulation used a single sample to estimate all parameters. As a result, the quality of estimation are different among the parameters. It is suspected that certain parameters have different optimal sample of other parameters. It suggests us to investigate more deeply the sample characteristic correspond to individual paramater. The estimation method of parameter individualy could be considered in the next research.

Acknowledge. This work was financed by Kopertis VII East Java, Ministry of Education and Culture, Republic of Indonesia on the Fundamental Research Project under contract No. 022/SP2H/P/K7/KM/2014.

References

[1] Aster, R. C, Borcher, B. and Clifford H. T., *Parameter Estimation and Inverse Problems*, Elseiver Academic Press, Burlington, MA, 2005.

[2] Bank, H.T, Arias, A.C. and Kappel, F., Parameter Selection Methods in Invers Problems Formulation, *Technical Report, CRSC, North Carolina State University*, Nopember (2010).

[3] Bank, H.T, Dediu, S. and Ernstberger, S.L., Sensitivity Function and Their Use in Inverse Problems, *J. Inverse Ill-posed Problem*, **15** (2007), 683-708.

[4] Bank, H.T, Dediu, S., Ernstberger, S.L. and Kappel, F, Generalized Sensitivity and Optimal Experimental Design, *J. Inverse Ill-posed Problem*, **18** (2010), 25-88.

[5] Bank, H.T, Holm, K. and Kappel, F, Comparation of Optimal Design Methods in Inverse Problems, *Technical Report TR10-11, CRSC, North Carolina State University, Inverse Problems, Submitted*, September (2010)

[6] Doicu, A., Trautmann, and Schreier, F., *Numerical Regularization for Atmospheric Invers Problems*,Springer-Verlag, Berlin Heidelberg, 2010.

[7] Goodwin, C.G. and Payne, R.L., *Dynamic System Identification: Experimental Design and Data Analysis*,Academic-Press, 1977.

[8] Hernadi, J., The Experimental Strategy for Parameter Estimation of Dynamic System Models, *Preprint*, (2011)

[9] Hernadi, J and and Kappel, F, Optimal Design of Experiment: Sensitivity and Parameter Identification, *Technical Report KF Univ-Graz, Austria*, Desember (2010)

[10] Muller, W.G., *Collecting Spatial Data: Optimum Design of Experiment for Random Fields,* Springer-Verlag, Berlin Heidelberg, 2007.

[11] Thomaseth, K. and Cobelli, C., Generalized Sensitivity Functions in Physiological System Identification, *Annals of Biomedical Engineering*, **27** (1999), 607-616.

[12] Ucinski, D., *Optimal Measurement Methods for Distributed Parameter System Identification,* CRC Press LLC, 2005.

[13] Vogel C.R., *Computational Methods for Inverse Problems,* SIAM, Frontier in Applied Mathematics FR34, Philadelphia, 2002.

Julan HERNADI: Department of Mathematics Education, Faculty of Training Teacher and Education, Muhammadiyah univ of Ponorogo.
E-mails: julan_hernadi@yahoo.com