Proceedings of IPC'02
4^{th} International Pipeline Conference
September 29-October 3, 2002, Calgary, Alberta, Canada

IPC2002-27145

A TIME SERIES APPROACH FOR PIPE NETWORK SIMULATION

André L. H. Costa
UFRJ, Escola de Química
Centro de Tecnologia, Bl. E, Ilha do Fundão
Rio de Janeiro –RJ - 21945-900 – BRAZIL
andrehc@uninet.com.br

José L. de Medeiros
UFRJ, Escola de Química
Centro de Tecnologia, Bl. E, Ilha do Fundão
Rio de Janeiro –RJ - 21945-900 – BRAZIL
jlm@eq.ufrj.br

Ofélia Q. F. Araújo
UFRJ, Escola de Química
Centro de Tecnologia, Bl. E, Ilha do Fundão
Rio de Janeiro –RJ – 21945-900 – BRAZIL
ofelia@eq.ufrj.br

ABSTRACT

We applied time series predicting tools for the simulation of the temporal behavior of large pipeline networks submitted to timely changing inputs. The inputs may consist of a set of specified flow rates at client or supply nodes, while the outputs are another set of nodal pressures and internal flow rates. According to the topology, size, age and history of the network, the continuous generation of phenomenological dynamic simulations may be impossible, imprecise or numerically expensive, demanding thus alternative approaches. Our methodology is particularly oriented to this kind of demand. From recorded network past data covering relevant history of inputs and selected outputs, ARX-MIMO predictors are built with identification methods and launched for continuous estimation of the network outputs one time step ahead. Results are precise enough for engineering, training and monitoring applications.

Keywords: Simulation, Pipeline Network, Time Series, ARX

INTRODUCTION

This paper presents the use of process identification techniques for developing models able to capture the dynamic behavior of pipe networks. The availability of a model is important to applications such as dynamic process simulations, process monitoring, control and staff training in the operation of a pipe network. According to the topology, size, age and history of the network, the continuous generation of phenomenological dynamic simulations may be impossible, imprecise or numerically expensive, demanding thus alternative and adaptive approaches. We applied time series predicting tools for the simulation of the temporal behavior of large pipeline networks submitted to timely changing inputs. The inputs may consist of a set of specified flow rates at client or supply nodes, while the outputs are another set of nodal pressures and internal flow rates. Our approach is only supported by time series data of the main process variables with no need for phenomenological knowledge of the system structure or its parameters (e.g. pipe connection diagram, pipe roughness, pipe diameters, etc.). In this work we considered the simulation of a hypothetical pipe network composed of several supply and demand sites interconnected, including the presence of loops in the structure, and adopting typical large pipeline diameters and lengths. A pseudo-stationary simulation algorithm was employed to generate the network response under the influence of time series of input data, which follow a temporal pattern corrupted by typical noise, thus emulating the dynamic behavior of the real process. The predictor is built based on an ARX model (auto-regressive with extra inputs) in a MIMO structure (multiple-input, multiple-output). We investigated two polices of mounting and using the predictor, which are defined by the specification of the lengths of the training and predicting phases. Results indicate that the simulator allows a good representation of the network behavior, predicting output values close to those generated by the rigorous model employed to represent the real system, and corroborated by an uncertainty analysis based on confidence limits.

NOMENCLATURE

$\underline{\underline{A}}(q)$: AutoRegressive filter with order n

$\underline{\underline{A}}_k$: constant matrix ($ny$ x ny)

$\underline{A}_0$, $\underline{A}_1$: vectors of pump parameters
$\underline{A}_2$, $\underline{A}_3$: vectors of pump parameters
$\underline{\underline{B}}(q)$: eXogenous input filter with order *m*
$\underline{\underline{B}}_k$: constant matrix (*ny x nu*)
$\underline{D}$: vector of tube diameters
E, E_P : numbers of tube edges and pump edges
$\underline{\underline{F}}$ selection matrix for assignment of specifications
$\underline{f}$: vector of Darcy friction factors
g : acceleration of gravity
$\underline{H}$: vector of pump heads
$\underline{K}$: vector of tube hydraulic parameters
$\underline{L}$: vector of tube lengths
$\underline{\underline{M}}$: network incidence matrix
N : number of network nodes
NT : number of time instants in the training phase
NI : number of time instants in the predictive phase
n,m : orders of AR and X filters in ARX model
nu,ny: lengths of input and output vectors in ARX model
$\underline{P}$: vector of nodal pressures
$\underline{Q}$: vector of tube flow rates
$\underline{Q}_U$: vector of upper bounds for pump flow rates
q : forward shift operator
$\underline{Re}$: vector of Reynolds numbers
$\underline{\underline{S}}$: selection matrix
t: time instant
$\underline{U}(t)$: process inputs at time instant *t*
$\underline{V}(t)$: process specifications at time instant *t*
$\underline{W}$: vector of node external flow rates
$\underline{X}(t)$: process variables at time instant *t*
$\underline{Y}(t)$: vector of outputs or measured process variables
$\hat{\underline{Y}}(t+1)$: predicted vector of outputs for instant *t+1*
$\underline{Z}$: node heights
$\underline{\alpha}$: vector of slacks for pump flow rates
$\underline{\beta}$: vector of slacks for pump flow rates
ε: pipe roughness
ρ, μ : fluid density and viscosity

1. GENERAL ASPECTS ON PIPELINE NETWORK SIMULATION

Large pipeline networks are becoming very important systems in modern human society, responding by the transport and distribution of potable and irrigation water, fuels, industrial raw materials like crude oil and natural gas, and several petrochemical and industrial fluids. Due to worldwide continuous expansion of pipeline network applications, they also exhibit trends related to gradually larger sizes; higher degree of topologic complexity in designs and increasing levels of service pressures and flow rates. In this context, management and logistics of large-scale fluid transportation companies depend crucially on routine tasks related to precise dynamic monitoring and forecasting of pipeline networks. These activities demand, on the other hand, computing resource to predict the network dynamic or stationary behavior. The dynamic simulator is a broader concept, as one stationary state of a system is merely a point in one of its admissible dynamic trajectories. Using a discrete time approach, a dynamic simulator is conceived as a black- or white-box specialized in reproducing the network state at time instant t provided enough information on its previous state t-1 and the corresponding inputs at t-1. The white-box approach, a phenomenological network simulator, is characterized by strict physical modeling based on nodal balance equations, tube momentum balances, pump operation curves etc. The white-box may also exhibit uncertainties in parameters or in some of its theoretical components. This is especially true as the system departs from its original condition by aging and/or by stressing action of exogenous agents. In this case, the model gradually loses its "whiteness" evolving towards what is called a gray-box, i.e. a phenomenological model containing non-negligible uncertainties in its constitution. Senile phenomenological models thus have to be "rejuvenated" by training sessions where its unknown components or parameters are re-estimated from records of the trajectory of the real system. The black-box, on the other side, does not contain any physical principle in its definition. It is, nevertheless, built with specialized resources for high adherence to past history of the system, such that its near future can be predicted if the set of inputs that will reach the system could be known. The black-box thus depends on training phases where its internal parameters are updated in order to capture the essential patterns of the system dynamics. If continuous training is not allowed, the black-box also suffers from gradual "aging effects" rapidly losing adherence to the system response. The state of the network can be invariably associated with a set of outputs (e.g. pressures at client nodes, flow rates at supply nodes, pump heads and internal or edge flow rates), which dynamically evolves responding to a set of inputs (e.g. pressures at supply nodes, flow rates at client nodes). The dynamic simulator – either white, gray or black box – is thus useful to predict: (i) new scenarios associated to changes in the characteristics of served (clients) or supply sites; (ii) dynamic effects resulting from the interaction of combined dynamic patterns of different inputs, in order to prevent dangerous situations on certain network sites; (iii) operational limits of the network; (iv) dynamic response of the network under the action of stochastic behavior of inputs, also for testing operational limits and safety levels; and (v) training of personnel. In this work, we consider the development of two strategies for simulation of pipeline networks: a white-box simulator and a black-box simulator. The white-box model is a simplified description based on: (i) incompressible newtonian fluid flow; (ii) pump modeling via cubic operational curves; (iii) no delays in the system response, meaning that the model is in fact a stationary description of the network, responding in pseudo-stationary mode to the time series of inputs. This model is able to emulate the real behavior of networks and was used for training the black-box model, which is a predicting tool with ARX-MIMO structure adopting $n=3$, $m=3$, $ny=8$, $nu=8$, where n and m are respectively the orders of the auto-regressive and input filters, ny and nu are respectively the sizes of the vectors $\underline{Y}$ of network outputs or measurements (pressures at client nodes) and $\underline{U}$ of network inputs (flow rates at client nodes).

2. PSEUDO-STATIONARY MODELING OF PIPELINE NETWORK WITH INCOMPRESSIBLE FLOW

A model is required to simulate a network plant in a simple way while retaining basic complexities as non-linear input/output relationships, multi-tube configuration and incidence of noises. Thus, a simple dynamic pipeline network is proposed, which generates no-delayed, pseudo-stationary responses to a selected group of inputs distributed in time according to certain stochastic patterns. The plant is represented by a direct graph whose E pre-oriented edges are pipelines connecting N nodes. There are Ep edges ($Ep \leq E$) with pumps, one pump per edge. The nodes have streams to/from the external environment (W). Process connectivity is described by an incidence matrix ($\underline{\underline{M}}$), such that $M(i,j) = (+/-)\ 1$ implies that edge j arrives at/leaves node i; otherwise $M(i,j) = 0$. The vector $\underline{X}(t)$ of process variables at time instant t is composed by N nodal pressures ($\underline{P}$), E edge (signed) flow rates ($\underline{Q}$), N external (signed) flow rates ($\underline{W}$), Ep pump heads ($\underline{H}$), Ep slack variables $\underline{\alpha}$ for lower bounding the flow rates in pump edges, and Ep slack variables $\underline{\beta}$ for upper bounding the flow rates in pump edges. The vector of Ep pumps is characterized by vectors $\underline{A}_0, \underline{A}_1, \underline{A}_2, \underline{A}_3$ containing each Ep parameters for the respective operational curves, and by vector $\underline{Q}_U$ containing Ep upper bounds of pump flow rates. There is also a selection matrix $\underline{\underline{S}}$ (size Ep by E) assigning edges to pump edges. Model pseudo-stationary relationships (Eq. (1)) are N nodal mass balances, E edge momentum balances (Eq. (2)), $2.Ep$ equations for bounding flow rates in pump edges (Eq. (3), (4)), and Ep operational equations for pumps (Eq. (5)). All rates are expressed in mass basis and SI units are used. The remaining N degrees of freedom are assigned to a vector $\underline{V}(t)$ of specifications – one specification per node – formed with selected nodal pressures and nodal flow rates. For instance, nodes that are merely junction points may be specified with $W=0$, while supply nodes with fixed operational patterns may be specified with a high pressure value. Vector $\underline{U}(t)$ of inputs is a sub-set of $\underline{V}(t)$ containing nu components. The inputs are typically flow rates at client nodes. Input signals are generated as time series by superposition of pseudo-random and gaussian noises over nominal values of the respective specifications as shown in Section 5. Specification values are assigned to vector $\underline{X}(t)$ by means of another selection matrix $\underline{\underline{F}}$ (size N by $N+E+N+3*Ep$) according to Eq. (6).Vector $\underline{Y}(t)$, containing ny outputs, can be any measured sub-set of $\underline{X}(t)$. In this work, the outputs were assigned to pressures at client nodes. Other entities are N node heights (Z), E edge Darcy friction factors ($\underline{f}$), lengths ($\underline{L}$), diameters ($\underline{D}$), tube hydraulic parameters ($\underline{K} = (8/\rho\pi^2)\underline{L}/\underline{D}^5$), Reynolds numbers ($Re$), where ρ, μ and g are density and viscosity of the fluid and the gravitational constant. Friction factors ($\underline{f}$) are predicted by Churchill equation [1] from ($\underline{Re}$). For each time instant t, with input and specification vectors $\underline{U}(t)$ and $\underline{V}(t)$, the set of states (and responses) $\underline{X}(t)$ (and $\underline{Y}(t)$) is obtained by numerical resolution of the network equations by a Newton-Raphson algorithm [1].

$$\underline{\underline{M}} \cdot \underline{Q} + \underline{W} = \underline{0} \tag{1}$$

$$\underline{\underline{M}}^t \cdot \underline{P} + \underline{f}.\underline{K}.\left|\underline{Q}\right|.\underline{Q} + \rho \cdot g \cdot \underline{\underline{M}}^t \cdot \underline{Z} - \rho.g.\underline{\underline{S}}^t.\underline{H} = \underline{0} \tag{2}$$

$$\underline{\underline{S}}.\underline{Q} - \underline{\alpha}.\underline{\alpha} = \underline{0} \tag{3}$$

$$\underline{\underline{S}}.\underline{Q} + \underline{\beta}.\underline{\beta} - \underline{Q}_U = \underline{0} \tag{4}$$

$$\underline{H} - \underline{A}_0 - \underline{A}_1.(\underline{\underline{S}}.\underline{Q}) - \underline{A}_2.(\underline{\underline{S}}.\underline{Q}).(\underline{\underline{S}}.\underline{Q}) - \underline{A}_3.(\underline{\underline{S}}.\underline{Q}).(\underline{\underline{S}}.\underline{Q}).(\underline{\underline{S}}.\underline{Q}) = \underline{0} \tag{5}$$

$$\underline{\underline{F}}.\underline{X}(t) - \underline{V}(t) = \underline{0} \tag{6}$$

3. PIPELINE NETWORK SIMULATION BY ARX-MIMO PREDICTORS

In a real application, recorded measured data covering network inputs and selected outputs are used for training ARX-MIMO predictors [2]. In the present case, the network time series, generated as shown in Section 2, will be used for training the predictors. The ARX-MIMO predictor has the following form:

$$\hat{\underline{Y}}(t+1) = \underline{\underline{B}}(q)\underline{U}(t+1) + q\left[I - \underline{\underline{A}}(q)\right]\underline{Y}(t) \tag{7}$$

with $\underline{\underline{A}}(q) = 1 + \sum_{k=1}^{n} \underline{\underline{A}}_k q^{-k}(.)$ and $\underline{\underline{B}}(q) = \sum_{k=1}^{m} \underline{\underline{B}}_k q^{-k}(.)$

where $\hat{\underline{Y}}(t+1)$ is the predicted vector of outputs for instant $t+1$; q represents the forward shift operator; $\underline{\underline{A}}(q)$ and $\underline{\underline{B}}(q)$ are matrix filters – respectively the AutoRegressive filter with order *n* and the eXogenous input filter with order *m*, $\underline{A}_k$ and $\underline{B}_k$ are constant matrices with sizes (*ny x ny*) and (*ny x nu*), respectively. The predictor identification consists in estimating matrices $\underline{A}_k$ and $\underline{B}_k$ from recorded values of $\underline{Y}$ and $\underline{U}$ during a training window with *NT* time instants (the training phase). This is a linear estimation problem that can be conducted via standard techniques. The predictor is used subsequently for the next *NI* instants estimating process outputs (the predictive phase). Training and predictive phases are then repeated. In routine applications, *NI* can be three to five times larger than *NT*. The quality of the training and the degree of deterioration of the predictor, as it departs from the training window, can be measured through the semi-width of *99%* confidence intervals for correct responses in the training and predictive phases: narrow confidence intervals mean high adherence to the process outputs; and, vice-versa, large confidence intervals are symptoms of bad reproduction of the process. In general, the predictor is more adherent to the process in the training window. This adherence deteriorates as it departs from the training window and experiences input signals very different from the corresponding training set. The global performance of the predictor depends on: (i) the set of measurements or monitored responses (Y); (ii) the orders and sizes (*n,m,ny,nu,NT,NI*) that characterize the predictor; (iii) the intensity of noises acting on the input vector in the training window and outside it; (iv) the intensity of noise acting on the measurements. The use of *n=3, m=3, ny=8 and nu=8,* imply that there are *n*ny+m*nu=48* parameters for each response to be predicted, with 24 parameters of type *A* and other 24 of type *B* (Eq. (7)).

4. NUMERICAL RESULTS

The capability of the proposed approach to reproduce the dynamic behavior of a pipe network is presented by a numerical example. Initially, the dynamic simulation of a network is conducted through the mentioned pseudo-stationary scheme. The purpose of this simulation is to generate the time series data equivalent to the measured data, which would be available from the process instrumentation in a real application. The structure of the simulated network is shown in Fig. 1, where node 1 is the supply site and the others nodes are demand sites. All pipes are *45000*m long having different diameters according to Table 1. Pumps are not present in the interior of the network. The absolute roughness of pipes is $\varepsilon = 46 \cdot 10^{-6}$ m. Diameters (m) adopted for each pipe are respectively *{.61; .25; .25; .46; .15; .15; .41; .20; .25; .25; .15}*. Nodes, numbered 1 through 9, have elevations (in meters) given by *{200; 200; 190; 195; 190; 190; 195; 205; 210}*.

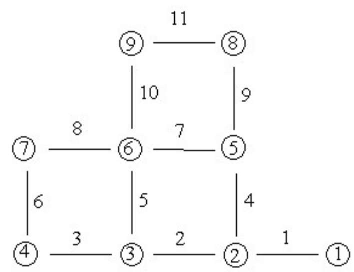
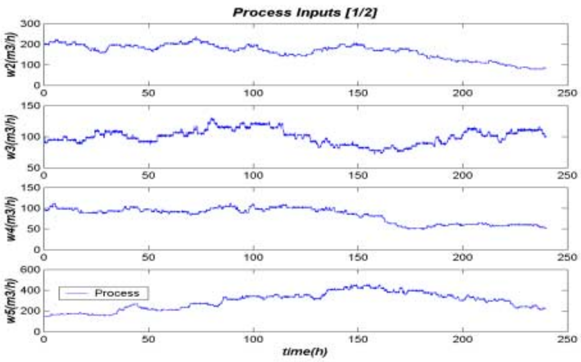


Figure1. Network Structure

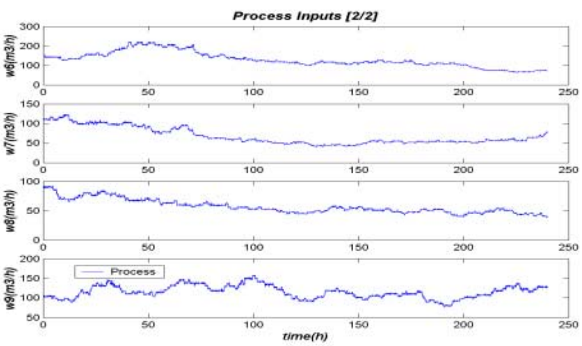
The fluid in the network is a diesel like oil with the following values of density and viscosity : $\rho = 833$ kg/m^3 and $\mu = 7.25 \cdot 10^{-3}$ Ns/m^2 = *8.7*cSt. The input variables of the example are the set of demand flow rates ($\underline{U}^T$*=[W2... W9]*) expressed in volumetric hourly basis (m3/h) and with no sign. The output (measurements) variables are the pressures at the respective nodes *(Y^T =[P2 ... P9])*. The vector of nine network specifications ($\underline{V}(t)$) is completed with the pressure at the supply node, which is kept constant (*P1=70 bar*). Measurement errors are not included in the analysis. The independent dynamic behavior of the input variables are defined by the following algorithm, applied to all time instants:

For $K = 2,...,9$:
Let $\delta_1 \in [0,1]$ and $\delta_2 \in [0,1]$ be two independent uniform random variables.
If $\delta_1 \le a$ then $usp_K \leftarrow usp_K + usp_K \cdot \Delta usp \cdot \eta_1$;
If $\delta_2 \le b$ then $u_K \leftarrow usp_K + usp_K \cdot \Delta u \cdot \eta_2$;
Otherwise $u_K \leftarrow usp_K$.

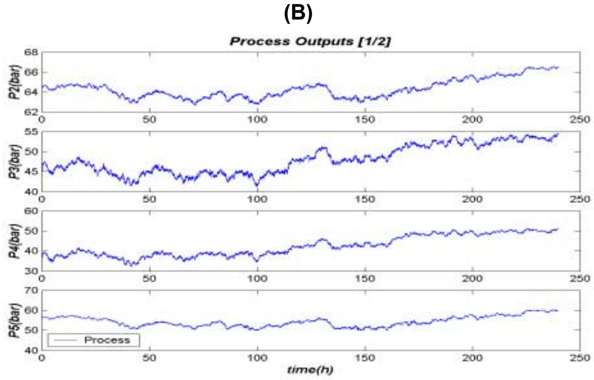
where $a = 0.02$, $b = 0.50$, $\Delta usp = 0.05$, $\Delta u = 0.01$, $\eta_1 \in \{-1,1\}$ is an uniform random variable and $\eta_2 \in \Re$ is a random variable with standard normal distribution. Symbol *usp* represents the value of an input variable in the previous time instant. The initial values of *usp* are $usp_2 = 200$ m^3/h, $usp_3 = 100$ m^3/h, $usp_4 = 100$ m^3/h, $usp_5 = 150$ m^3/h, $usp_6 = 150$ m^3/h, $usp_7 = 100$ m^3/h, $usp_8 = 90$ m^3/h and $usp_9 = 110$ m^3/h. This algorithm may represent two phenomena: a frequent and small fluctuation of input variables along the time (noise) and a larger modification of the variable value indicating changes in the operating conditions. The time span of the simulation is *240* h. Input/output data are sampled with an interval of *2* min, leading to *7200* sampling instants. Figures 2-A and 2-B show the time series of the eight input variables acting on the process. Figures 2-C and 2-D depict the time series of the eight process responses generated as explained in Section 2. In order to provide some measure about the durability of the predictor, it was identified and used, according to two polices: (i) Police *1* with *NT=1000* and



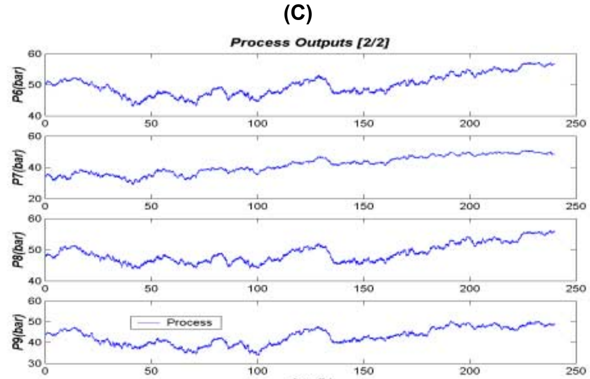
(A)



(B)



(C)

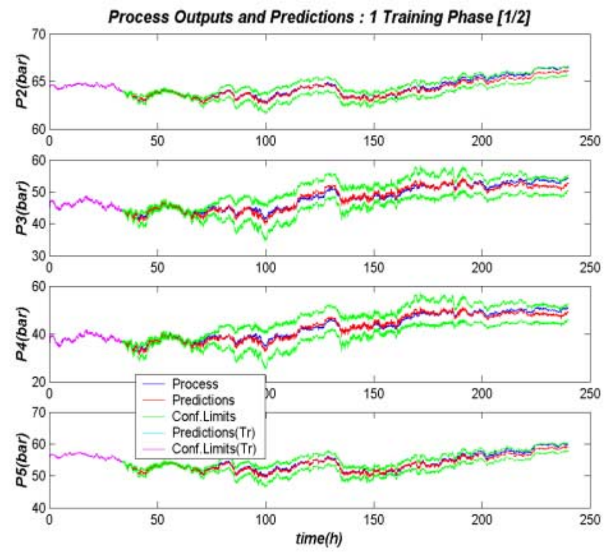


(D)

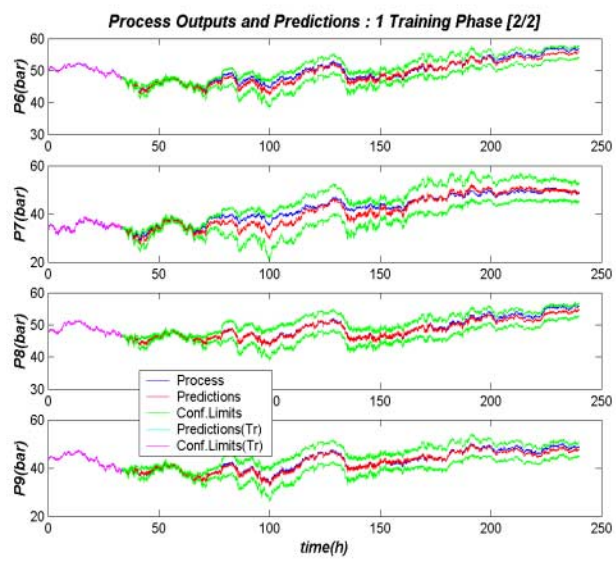
Figure 2: Process Inputs (Node Flow Rates) and Process Outputs (Node Pressures)

NI=6200 (predictor is adjusted using the first recorded *33.3* hours of the run and launched for the remaining *206.7* hours); (ii) Police *2* with *NT=1000* and *NI=2600* (predictor is adjusted using a record of *33.3* hours, launched for the next *86.7* hours; these steps are repeated until completion of the entire run).

Figures 3-A and 3-B depict the predictor performance according to Police *1*, while Fig. 4 exhibits the histogram of predicting errors associated with Police *1*. The time series of Police *2* are displayed in Figs. 5-A and 5-B. In Figs. 3 and 5, the *99%* confidence intervals are colored with magenta and green, respectively, in the training and predicting phases.



(A)



(B)

Figure 3: Process Outputs, Predictions and 99% Confidence Limits for Correct Values Using One Training Phase (Tr : Training Window) [Police 1]

Figures 6-A and 6-B present, respectively, the magnitude of the *48* ARX parameters and the corresponding standard deviations for each predicted response in the second training phase of Police *2*. It can be seen that the first *24* parameters (i.e. belonging to the AR filter) exhibit larger magnitudes when compared with the last *24* parameters (i.e. belonging to the X filter). A partial reason for this may be attributed to the larger magnitudes of the input signals relatively to the outputs. Another reason may be located in a certain preference of the ARX structure for using the auto-regressive processor. Another visible fact is that the relative importance of a parameter in the ARX structure decays with the order of the matrix term where it is situated. Figure 7 is another representation of the degree of uncertainty of each estimated ARX parameter by means of paired projections of the *99%* global parameter confidence domain: the lower the uncertainty associated to a given pair of parameters, the shrunken its ellipsoidal domain.

Finally, it is evident by comparing the results of Polices *1* and *2*, that the predictor is a more dependable tool when it is refreshed more frequently, i.e. it obviously suffers aging effects since the process may be subject to very different class of inputs as time passes.

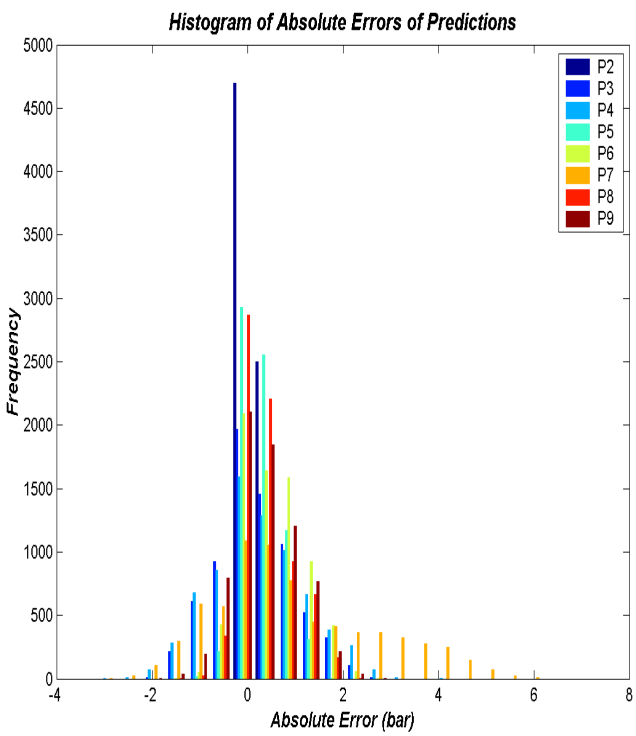
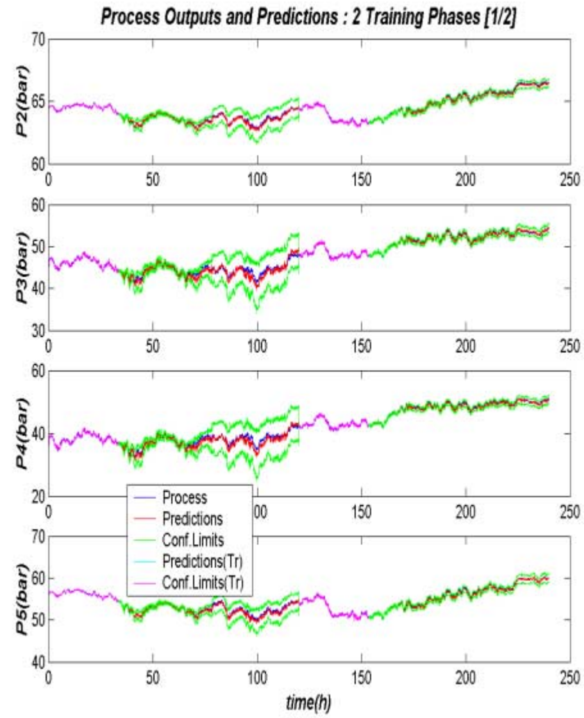
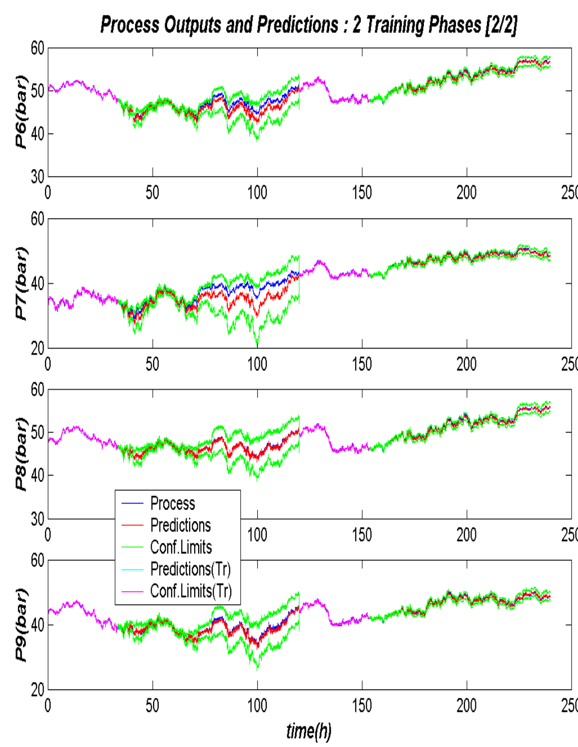


Figure 4: Histogram of Absolute Errors of Predictions (Node Pressures) Using One Training Phase (See Figure 2) [Police 1]



(A)



(B)

Figure 5: Process Outputs, Predictions and 99% Confidence Limits for Correct Values Using Two Training Phases (Tr: Training Window) [Police 2]

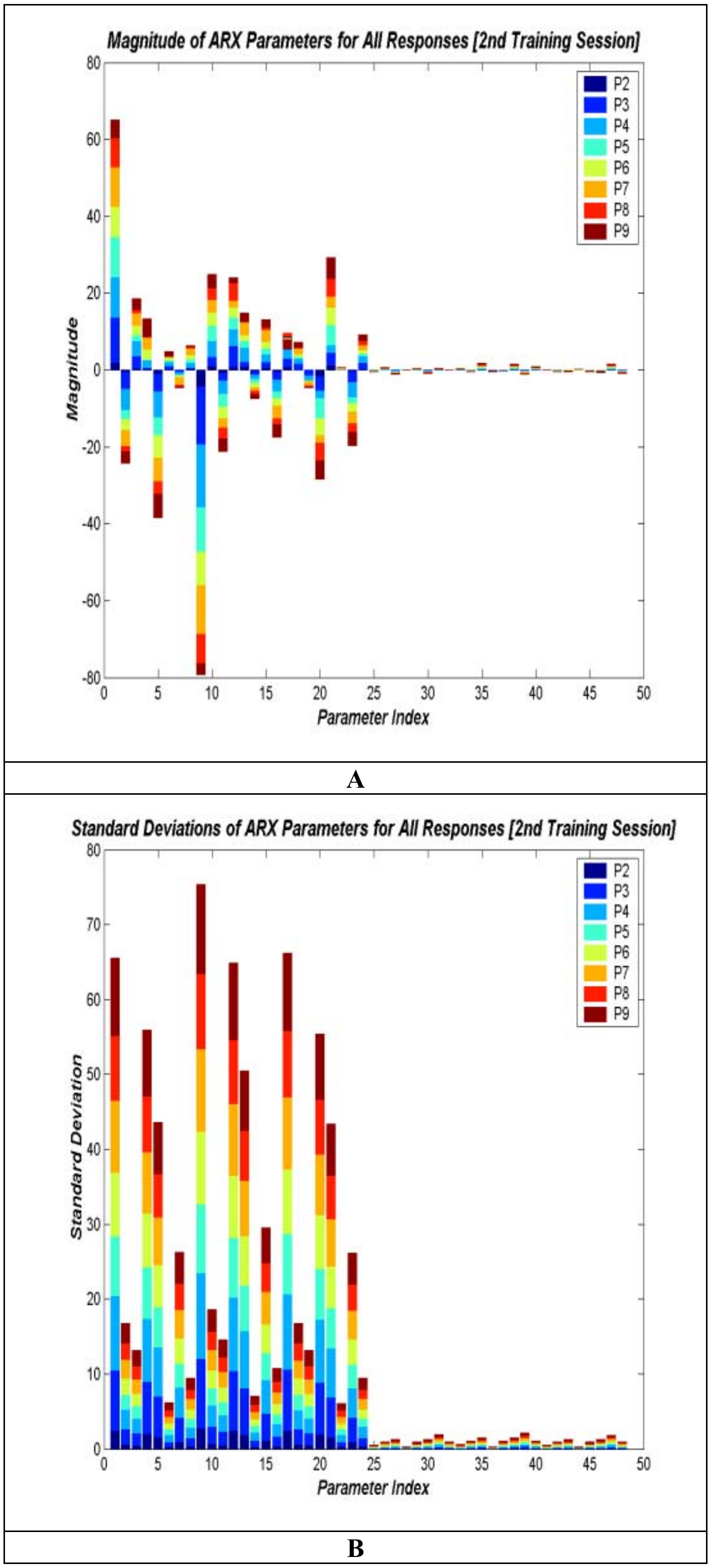


Figure 6: (A) Magnitude of ARX Parameters (B) Standard Deviations (%) of ARX Parameters (Using Two Training Phases [Police 2])

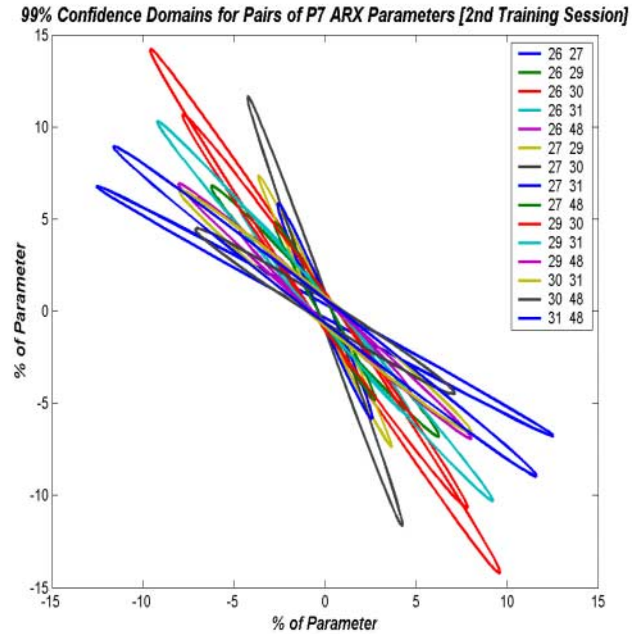


Figure 7: 99% Confidence Domains for Some Pairs of ARX Parameters for Predicting P7 Using Two Training Phases (Values for the 2nd Training [Police 2])

5. CONCLUDING REMARKS

This work proposes a black box approach for the reproduction of the dynamic behavior of large pipeline networks, based on system identification techniques. Particularly we used an ARX-MIMO *[3,3,8x8]* predictor to simulate the temporal behavior of a network subjected to stochastic patterns of flow rates at client nodes. The process response was emulated by a pseudo-stationary network simulation model, for incompressible, single-fluid flow, responding instantaneously to time series of inputs. The methodology was able to predict the system behavior within reasonable accuracy levels. The adherence of the method could be improved adopting more frequent predictor training sessions.

ACKNOWLEDGMENTS

Research Assistantship from CNPq (National Council for Scientific and Technological Development and Research) and Grant from FINEP/CTPETRO and ANP (National Petroleum Agency) are gratefully acknowledged.

REFERENCES

[1] Costa, A. L. H., 1997, “Steady-State Modeling and Simulation of Pipe Networks”, M.Sc. thesis (in Portuguese), Escola de Química, UFRJ, Rio de Janeiro.

[2] Box, G., Jenkins, G., Reinsel, C., 1994, *Time Series Analysis, Forecast and Control*, Prentice Hall, New York.

