Fault adaptive control: towards robust operation of autonomous systems

Gautam Biswas, Eric-J. Manders, Sherif Abdelwahed, Jian Wu, and John Ramirez
Department of Electrical Engineering and Computer Science,
and the Institute for Software Integrated Systems,
Vanderbilt University, P.O. Box 1824 Station B, Nashville, TN 37235-1592 USA.
Phone:+1 615 322-0732,
Email: {gautam.biswas}@vanderbilt.edu

Abstract— **Fault adaptive behavior is an increasingly important aspect for achieving autonomy in complex controlled physical systems. We describe a model-based approach for achieving this goal, termed Fault Adaptive Control Technology (FACT). FACT system models are constructed as hybrid bond graph models that capture both discrete and continuous behavior of the physical system. A runtime environment supports combined qualitative and quantitative fault detection, isolation and identification. Fault adaptivity is achieved through decision theoretic control approaches. We apply the approach to simulations of a real physical system testbed, a Water Recovery System that is a component of an Advanced Life Support system for extended duration human space missions. The effectiveness of the approach is illustrated in several fault scenarios.**

I. Introduction

The increasing complexity of technical systems and their use in safety-critical applications has imposed strict requirements on their reliability, robustness, and availability. This has led to renewed interest in testing, maintenance, and automated diagnosis of faults in these systems. Model-based approaches have become an increasingly attractive option as the field has matured and they include a wide range of techniques from traditional control systems approaches [1] to the use of artificial intelligence techniques [2].

In conventional diagnosis applications, the objective is to inform a human operator about the fault, through the generation of an 'alarm', and further determine its origin (fault isolation) and size (fault identification). An emerging application area is that of fault adaptivity, where the system autonomously responds to a fault in order to mitigate the impact of the fault on the operation of the system and to preserve the functionality as best as possible. Such an approach may combine adaptive tuning of controller parameters and decision making schemes that reconfigure the plant operation to compensate for, or eliminate altogether, the fault effects. In the area of embedded systems application, where the computational system is tightly coupled with the physical plant, fault adaptive control is becoming a critical capability. Applications areas include unmanned autonomous vehicles, and mission critical subsystems of large systems.

This paper describes a solution to this problem called *Fault Adaptive Control Technology* (FACT) [3]. FACT is designed using model-integrated computing techniques [4], and is supported by a tool set intended to allow a designer to explore performance of a fault adaptive controller on different fault scenarios. The tool suite comprises: (*i*) an environment for building dynamic models of the physical plant, its interface hardware that includes sensors and actuators, and the controller; (*ii*) a simulation environment based on the plant, interface, and controller models that allows for simulating nominal and faulty system data, and (*iii*) a computational environment and run-time support for FDI^2 and fault adaptive control as an embedded system application.

System behavior, including both discrete and continuous dynamics, is modeled using the Hybrid Bond Graph (HBG) modeling paradigm [5]. A hybrid observer scheme, an extended Kalman filter (EKF) combined with hybrid automata, is employed to track system behavior and compute the residual vector. On-line FDI is a hybrid extension of the TRANSCEND approach, a combined qualitative and qualitative model-based fault isolation scheme [6], [3], [7]. Fault adaptivity is achieved through decision theoretic model-predictive control.

This paper describes the physical system modeling scheme, the FDI^2 approach and fault adaptive capabilities of FACT. The approach is applied to a simulated version of an actual Water Recovery System, a key element in an Advanced Life Support (ALS) system for next generation manned space missions [8]. Section 2 describes the physical system, and its model development for diagnosis and control. Section 3 describes the FACT run-time architecture and operation. Section 4 describes experiments and results, and section 5 presents conclusions.

II. The Advanced Life Support – Water Recovery System

The support of human life in the hostile environment of space for extended duration missions is a complex challenge. Constraining the use of consumable resources requires an artificial environment that maximizes reuse of these resources. Such an environment will depend on complex technical systems that contain or interact with biological and chemical processes [8]. An *Advanced Life Support* system (ALS) must exhibit a high level of autonomy, so as not to detract from the mission specific tasks of the crew. For long duration missions this includes the ability to adapt to changing mission objectives and crew configurations, handle planned downtime of subsystems for maintenance operations, and respond to unplanned events such as faults in a system component.

A Water Recover System (WRS) system is a key element in any ALS, and a WRS testbed was designed and built at NASA Johnson Space Center (JSC) [9], [10]. The design includes four main components that involves biological, physical, and chemical processes in four main subsystems. Waste water enters the Biological Water Processor (BWP) that removes organic matter and ammonia from the water. The effluent of the BWP is fed into the Reverse Osmosis (RO) subsystem that removes inorganic matter using a membrane system. The RO can typically clean about 85% of its input waste water. The remaining 15%, a concentrated brine, is passed to the Air Evaporation System (AES), which recovers additional water from the brine through heat exchange mechanisms. A Post Processing System (PPS), treats the combined output of the RO and the AES to remove trace impurities and generate potable water. In this paper, we discuss the RO subsystem of the WRS.

A. The Reverse Osmosis (RO) sub-system

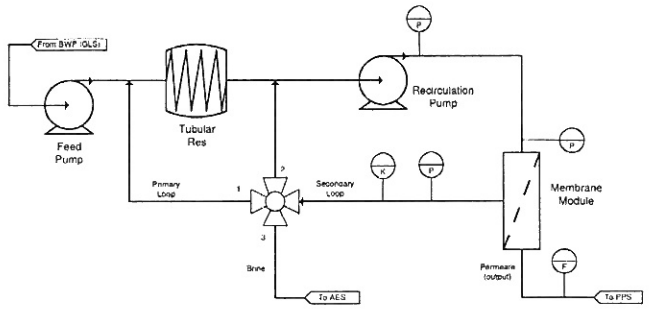

Fig. 1. High level process engineering diagram of the reverse osmosis (RO) system.

Fig. 1 shows the process engineering diagram with the primary components of the RO system [10], [8]. The system has one input, the output of the BWP, and two outputs, to the PPS and AES respectively. The key functional component of the system is a membrane [10]. The basic operation is to push input water at high pressure through the membrane. Clean water (permeate) leaves the system, and the remaining water, is recirculated in a feedback loop. As a result, the concentration of impurities in the recirculating water increases with time. The system cycles through three operating modes, which are determined by the position of a multi-position valve. The feed pump, which is always on, pulls effluent from the BWP and creates a flow into a tubular reservoir (coil). In the *primary* mode the input flow into the system is mixed with the recirculating water (recirculation loop). The recirculation pump boosts the pressure. A transition to the *secondary* mode occurs after a predetermined time interval. In secondary mode the recirculating fluid is directly fed back to the membrane in a smaller loop to increase flow rate and maintain sufficient flow through the membrane, whose resistance to flow increases as it accumulates dirt over time. The outflow of clean water from the loop causes an increase in brine concentration in the water remaining in the loop, and at a predetermined point that corresponds to 85% of volumetric recovery of water, a transition is made to the *purge* mode where the recirculation pump is turned off, and concentrated brine is pushed out to the AES subsystem. Following the purge operation, the system goes back to primary mode. A complete cycle (modes 1 through 3) takes approximately four hours.

The testbed has been extensively instrumented. Fig. 1 shows the five measured variables that are used for diagnosis in the current work: (*i*) the pressure immediately after the re-circulation pump, P_{pump}, (*ii*) the pressure of the permeate at the membrane, P_{memb}, (*iii*) the pressure of the liquid in the return path of the recirculation loop, P_{back}, (*iv*) the flow of the effluent, F_{perm}, and (*v*) the conductivity of liquid in the return path of the recirculation loop, K.

B. Modeling for diagnosis and control

Fig. 2 illustrates the HBG model created for this system, and consists of three main areas corresponding to three physical domains. Given the pump-fluid system, the *mechanical* and *hydraulic* domains are the primary energy domains that define the flow behavior in the system. However, to take into account the effects of impurities in the water on the flow process, and the fact that these impurities are time-varying, we explicitly model the fluid *conductivity* domain and its interactions with the flow process using bond graph elements. The energy interaction between the mechanical and the hydraulic domains is governed by the pump characteristics, which in our simplified models of the pump are represented by the pump efficiency. Depending on the pump type, this efficiency maps onto transformer and gyrator components in the bond graph modeling paradigm.

The model includes three switches (M1, M2, P) that correspond to the three modes of operation respectively. The control signals associated with each switch activate/deactivate the switching junctions in the model. Fig. 2 indicates the transition conditions for each switching junction as a logical expression of the switch values. In this model, all on/off transitions are inverse relations, and only the 'on' transition condition is specified.

The model captures the interaction between the hydraulic and conductivity domains in the bond graph using modulating signals. As a simplification, interaction between the mechanical and conductivity domains is abstracted away. The model describes the relevant behavior for diagnosis and control, rather than the physical structure of the system. Consider the feedback loops in the actual system, as compared to the modeling of feedback. While the physical system has two fluid loops (primary and secondary loops), the model has only one feedback loop. The effect of switching from primary to secondary mode is captured by changing the conductivity domain dynamics. using the modulation functions f_2 and f_3. The time varying nature of the membrane resistance is realized through the modulation function f_4.

To estimate the parameter values for the model, we determined a representative nominal behavior from the test-bed data, which is sampled time at 5 (min) intervals. Parts of the RO system were isolated, equations describing the parameter

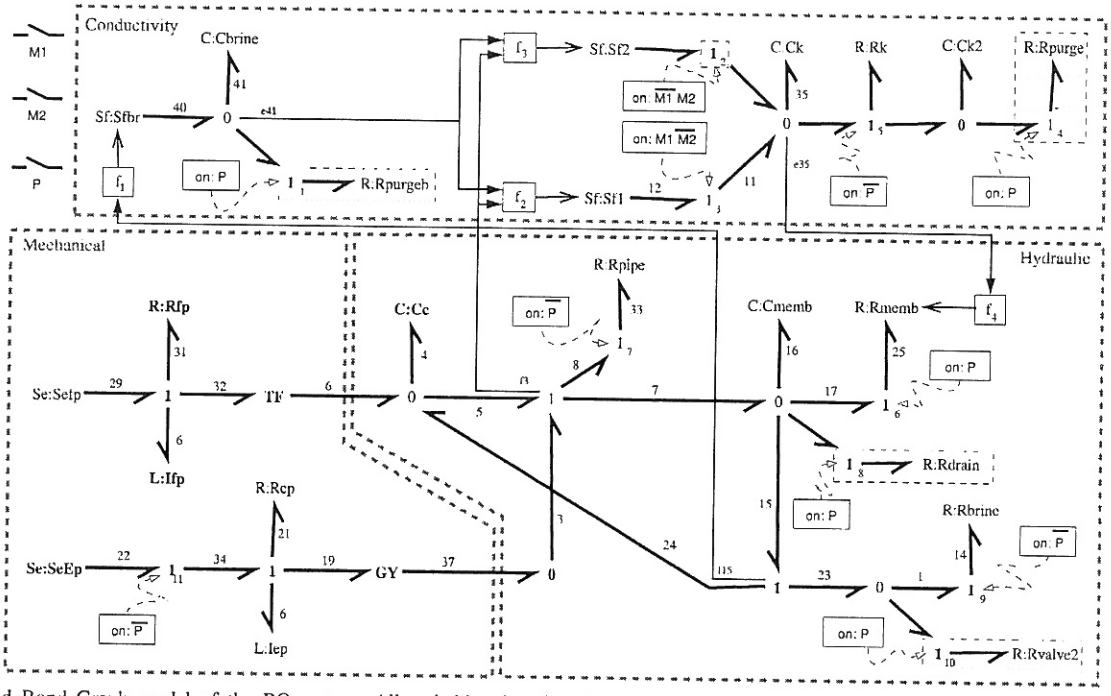

Fig. 2. Hybrid Bond Graph model of the RO system. All switching junctions are numbered, and shown with their corresponding transition conditions. Components that are not active in primary mode are shaded gray.

and measured value relations were derived from the bond graph, and parameters for the equations were estimated using least squares estimation. From this, bond graph component parameters were determined.

III. Model-based fault adaptive control

Fig. 3 shows the runtime system architecture, which is configured for a specific application based on the system model through a model-interpretation step. The active state model (ASM) is a dynamic component that maintains the current model of the system at run-time. It has two components: (*i*) the *structural model* that contains the HBG and its configuration of the current dynamics using the control signals and the output of the mode estimation unit, and (*ii*) the *model parameters*, a data structure that contains the current value of all model parameters. Either mode changes or identified faults may result in an update of model parameter values. From the ASM, the system generates multiple model representations that are used by run-time analysis components. (*i*) The extended Kalman filter in the hybrid observer and the optimization algorithm in the parameter estimation module for fault identification use a state-space model, (*i*) the qualitative fault diagnosis algorithms use a Temporal Causal Graph (TCG) model, and (*i*) the utility based controller uses a discrete time model representation.

We describe the FDI^2 and fault adaptive control components of the architecture in detail.

A. Transcend: *Fault detection, isolation and identification*

Transcend implements an FDI^2 scheme that can be interpreted as the typical FDI scheme consisting of residual generation and evaluation. The Transcend approach is based on detection and analysis of the system dynamics immediately

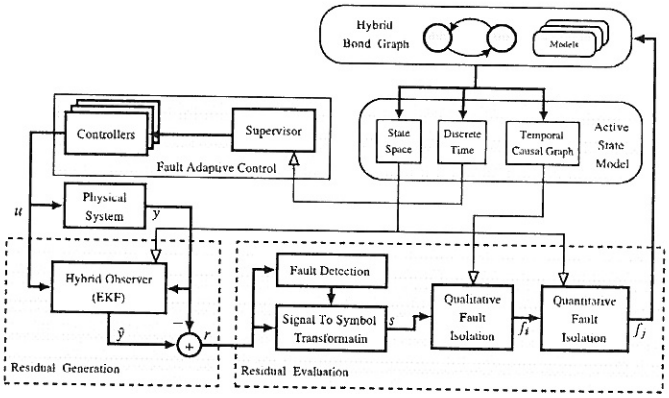

Fig. 3. FACT runtime system architecture.

after the time point of fault occurrence [7]. The hybrid observer includes an extended Kalman filter (EKF) [11], and a mode estimation component. The estimated state is used to determine any autonomous mode changes, and mode changes result in a adaptive update of the observer scheme. The estimated output variables are computed to allow the computation of a numerical residual as the difference between observed and estimated outputs. The residual signal is evaluated in a combined quantitative and qualitative analysis scheme.

Fault detection can be decoupled from the symbol generation for the qualitative fault isolation scheme. In this case sophisticated statistical signal processing techniques can be used to detect faults even when the dynamic response of the system results in a residual with a low signal-to-noise ratio. Alternatively, when fault detection is realized as a threshold on

individual residual signals, detection coincides with generating the first non-zero symbol. The signal to symbol transformation component is realized as a filter bank, where each symbol is realized using FIR filters whose outputs are quantized.

Qualitative model-based fault isolation, the core of the TRANSCEND approach, is a hypothesize-and-test scheme. Hypothesis generation is coupled with branching behavior to explore possible new modes that the system may have reached after the fault has occurred. The hypotheses for each fault candidate are matched against incoming symbolic measurement values in a TRANSCEND specific *progressive monitoring* scheme. When a match fails, the scheme evaluates if the observed behavior can be explained in a different mode of operation, and whether this warrants that the fault candidate should not be eliminated.

The fault isolation scheme is further enhanced with a quantitative parameter estimation step. In the current configuration, parameter estimation is initiated when qualitative fault isolation cannot reduce the candidate set any further. The parameter estimation scheme allows further fault isolation of those candidates that could not be resolved based on qualitative analysis. Because the changed parameter value is in fact estimated numerically, this also indicates the fault size (fault identification) [6].

B. Decision Theoretic Control

An on-line adaptive control mechanism implements a resource management scheme using a decision-theoretic controller based on a multi-attribute utility function that models system performance: $V(p) = \sum_i V_i(p_i)$, where each V_i corresponds to a value function associated with performance parameter, p_i. The parameters, p_i, can be continuous or discrete-valued, and they are derived from the system state variables, i.e., $p_i(t) = P_i(x(t))$. The value functions currently defined in our FACT paradigm are simple weighted functions of the form $V_i(p_i) = w_i * p_i$, where the weights take on values in the interval $[-1, 1]$, and represent the importance of the parameter in the overall operation of the system. The supervisory controller uses the active state model to predict possible behaviors corresponding to different action sequences for a finite forward time horizon, and then selects the action (i.e., control input) that maximizes the utility function. This process is then repeated for the next time step, and so on. Since the optimizing function operates on the current system model, the optimizing controller is fault adaptive, since it tailors its decision making to the model with the fault once the FDI[2] system has isolated and identified the fault parameter.

IV. EXPERIMENTS ON THE RO SYSTEM

We evaluate the FACT approach in several simulation experiments on the RO system. In each case we introduce faults in a bond graph component parameter, where a fault is a discrete change in the parameter value (abrupt fault profile). We add Gaussian noise to each measurement signal such that the noise power is 2% of the signal power. Aspects of detection sensitivity are not the focus of these experiments, and the fault size is chosen such that the probability of detection equals 1, and the fault can be completely isolated and identified. We describe a particular fault scenario in detail, and provide a summary overview of the results for other faults in the system.

A. Fault detection, isolation and identification

A decrease in the efficiency of the recirculating pump (effort pump) is modeled by a decrease in the value of a bond graph component parameter of the pump, the gyrator, GY. We indicate this fault scenario as GY^-. Fig. 4 shows the simulated plant data including the controller signals and the output of the observer, and the computed residual signals. The fault is introduced in the second operating cycle, while the system is in the primary operating mode.

Table I shows the fault isolation and identification results for the system for 5 fault scenarios. A scenario is described as the steps in the qualitative fault isolation sequence N), where step corresponds to an event when a new non-zero symbol is generated for a particular measurement (sensor), reflecting a qualitative change in the transient dynamics (symbols) of the fault response. The result of the step is shown as the set of candidates remaining. The '+' and '−' symbols correspond to an observed positive (negative) value in the magnitude field of the residual signal, and an positive or negative slope value in the slope field of the signal. A '0' implies no change, and a '·' value implies an unknown value. Step 0 is defined as the time when the fault is first detected, i.e., the first statistically significant non-zero residual is observed. The time at which this event occurs is marked as the difference between the current time and the time of fault occurrence, $t - t_f$. $t - t_f$ measures the delay in fault detection after the actual fault occurrence. The final entry is the result of the quantitative parameter estimation (fault identification) for the scenario. We discuss the GY scenario in detail. We discuss the GY^- scenario in detail. The predicted signatures for this scenario are not shown but may be found in [12].

At step 0 we observe that the residual for the pump pressure deviates in the negative direction ($P_{pump} = (-, \cdot)$). The hypothesis generation step that is triggered by this event results in twelve hypothesized fault candidates. The next observed change, the symbolic event at step 1, occurs when the flow-rate through the membrane also shows a negative deviation ($F_{perm} = (-, \cdot)$). Fault hypotheses whose signatures are not consistent with this (symbolic) residual value are dropped, and the fault set is reduced to seven hypotheses. At step 2, the deviation in the back flow pressure, P_{back}, becomes significant, and at step 3 the deviation in the conductivity, K, becomes significant also. When symbolic information can not discriminate between the remaining fault hypotheses, qualitative fault isolation terminates. Three candidates, GY^-, I_{ep}^+, and R_{ep}^+ remain after the qualitative fault isolation completes. In addition to the actual (true) fault candidate, these correspond to an increase in pump inertia (unlikely fault), and an increase in the pump resistance (e.g., increase in friction in the pump bearings), respectively.

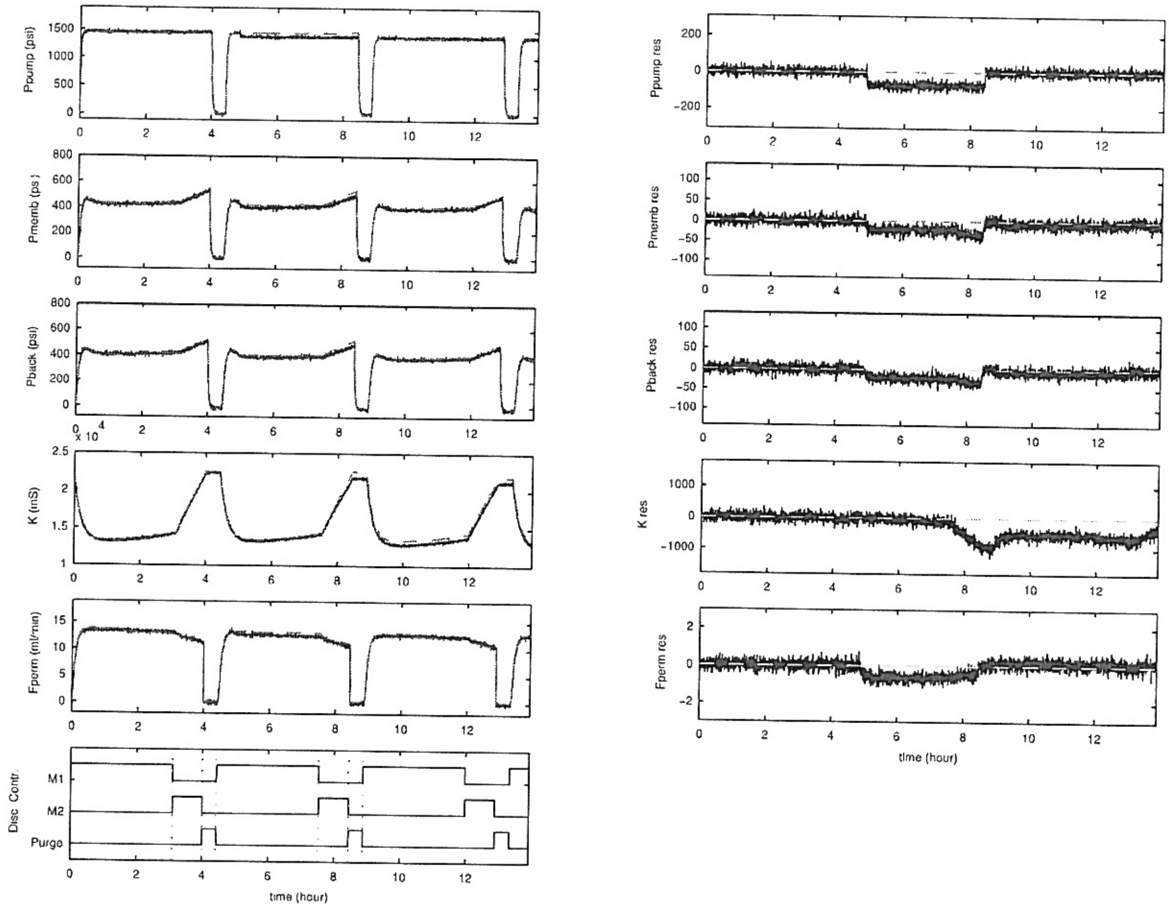

Fig. 4. Simulation data (l) and residual (r) for an abrupt fault GY^- with fault size 5%, occurring at $t_f = 17500$ (s) $= 5$ ($hours$) (primary mode in the second cycle).

Fault	$t - t_f$	N	Sensor	Symbolic	Candidate set + parameter estimation
GY^-, 5% t_f : 17500	200	0	P_{pump}	$(-,\cdot)$	C_c^+, C_{memb}^+, I_{fp}^+, I_{ep}^+, R_{brine}^-, TF^+, R_{pipe}^-, R_{memb}^-, C_k^+, R_{fp}^+, R_{ep}^+, GY^-
	880	1	F_{perm}	$(-,\cdot)$	I_{fp}^+, I_{ep}^+, R_{brine}^-, TF^+, R_{fp}^+, R_{ep}^+, GY^-
	1240	2	P_{back}	$(-,\cdot)$	I_{ep}^+, R_{brine}^-, R_{ep}^+, GY^-
	1960	3	K	$(-,\cdot)$	I_{ep}^+, R_{ep}^+, GY^-
					parameter estimation: GY^- changed by 0.934
R_{memb}^+, 5% t_f : 20000	800	0	F_{perm}	$(-,\cdot)$	C_c^+, C_{memb}^+, I_{fp}^+, I_{ep}^-, R_{brine}^-, TF^+, R_{pipe}^-, R_{memb}^+, R_{fp}^+, R_{ep}^-, GY^-
	7200	1	P_{back}	$(+,\cdot)$	I_{fp}^-, TF^+, R_{pipe}^-, R_{memb}^+, R_{fp}^+, R_{ep}^-
	8280	2	P_{memb}	$(+,\cdot)$	R_{pipe}^-, R_{memb}^+, R_{ep}^-
					parameter estimation: R_{memb}^+ changed by 1.042
R_{ep}^+, 35% t_f : 20000	88	0	P_{pump}	$(-,\cdot)$	C_c^+, C_{memb}^+ I_{fp}^+, I_{ep}^+, R_{brine}^-, TF^+, R_{pipe}^-, R_{memb}^-, C_k^+, R_{fp}^+, R_{ep}^+, GY
	640	1	F_{perm}	$(-,\cdot)$	I_{fp}^+, I_{ep}^+, R_{brine}^-, TF^+, R_{fp}^+, R_{ep}^+, GY^-
	720	2	P_{back}	$(-,\cdot)$	I_{ep}^+, R_{brine}^-, R_{ep}^+, GY^-
	960	3	P_{pump}	$(-,-)$	R_{brine}^-, R_{ep}^+
	4640	4	K	$(-,\cdot)$	R_{ep}^+
					parameter estimation: R_{ep}^+ changed by 0.374
R_{pipe}^+, 15% t_f : 18000	640	0	P_{back}	$(-,\cdot)$	C_c^-, C_{memb}^- I_{fp}^-, I_{ep}^+, R_{brine}^-, TF^+, R_{pipe}^+, R_{memb}^+, C_k^+, R_{fp}^-, R_{ep}^+, GY^+
	800	1	P_{memb}	$(-,\cdot)$	R_{brine}^-, TF^+, R_{pipe}^+, R_{ep}^+,
					parameter estimation: R_{pipe}^+ changed by 1.134
C_{memb}^-, 10% t_f : 19600	360	0	F_{perm}	$(-,\cdot)$	C_c^-, C_{memb} I_{fp}, I_{ep}^+, R_{brine}^+, TF^+, R_{pipe}^+, R_{memb}, C_k^+, R_{fp}, R_{ep}^+, GY^+
	480	1	P_{back}	$(-,\cdot)$	C_{memb}^- R_{brine}^+, TF^+, GY^+
	8680	2	P_{memb}	$(-,\cdot)$	C_{memb}^- R_{brine}^+, GY^+
					parameter estimation: C_{memb}^- changed by 0.856

TABLE I

COMPREHENSIVE DIAGNOSIS RESULTS FOR THE RO SYSTEM.

On completion of the parameter estimation, the hybrid observer is updated with the new parameter values, and continues to track the new behavior, where the known changed system behavior becomes the nominal behavior.

B. *Fault adaptive control for fault GY^-*

Fig. 5 shows the behavior of the system under on-line fault adaptive control as a result of the fault in the effort pump. The fault again occurs in the primary mode in the second cycle. The on-line controller compensates for the fault by changing the mode switching pattern, and keeping the system in primary mode for a longer time in each cycle. This is clearly visible when the system completes the third cycle, and the increase in cycle is easily noticeable. The overall average utility after the occurrence of the fault was only .93% less than the utility in the non-faulty situation. Details of this experiment can be found in [13]

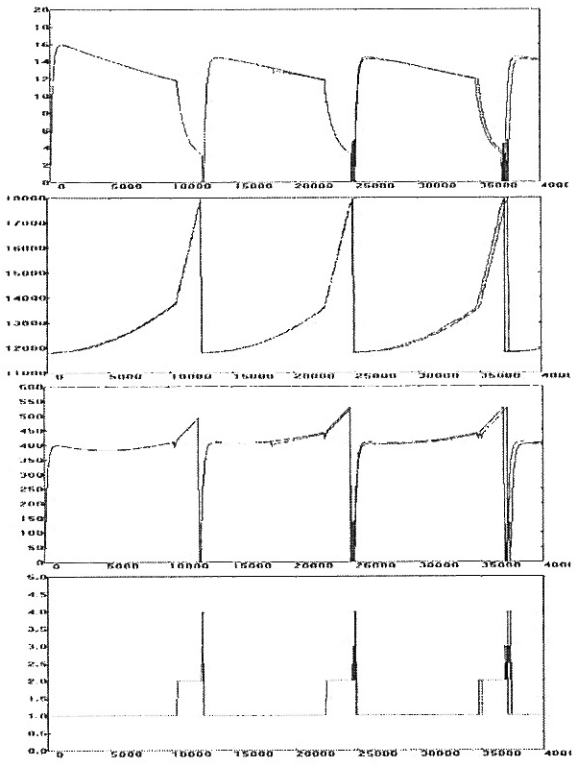

Fig. 5. On-line decision theoretic control for fault GY^-.

V. Discussion and Conclusions

This paper has described the FACT approach for developing model-based FDI[2] and fault adaptive control systems. The FACT implementation consists of a modeling paradigm and a run-time system that supports behavior tracking through observer schemes, fault detection, model-based fault isolation and fault identification, and model-predictive control. The scheme has been applied to the reverse osmosis component of a water recovery system.

In ongoing work we are applying the approach to other components of the WRS system, and developing multilevel control schemes for the integrated WRS system. The objective is to develop effective resource management schemes that adapt the system operation in response to fault occurrences, thereby improving the capabilities for autonomous operation of mission critical subsystems such as an ALS.

Acknowledgements

This work was supported in part through grants from the NASA-IS program (Contract number: NAS2-37143), DARPA SEC program (Contract number: F30602-96-2-0227), and the NASA-ALS program (Contract number: NCC 9-159). We acknowledge the help provided by Gabor Karsai and Gyula Simon in the development of the FACT architecture.

References

[1] P. M. Frank, S. X. Ding, and B. Köppen-Seliger, "Current developments in the theory of FDI," in *Proc. 4th IFAC Symp Fault Detection Supervision Safety Technical Processes*, Budapest, Hungary, 2000, pp. 16–27.

[2] P. M. Frank and B. Köppen-Seliger, "New developments using AI in fault diagnosis," *Engineering Applications of Artificial Intelligence*, vol. 10, no. 1, pp. 3–14, 1997.

[3] G. Karsai, G. Biswas, T. Pasternak, S. Narasimhan, G. Peceli, G. Simon, and T. Kovacshazy, "Towards fault-adaptive control of complex dynamical systems," in *Software-Enabled Control – Information Technology for Dynamical Systems*, T. Samad and G. Balas, Eds. NJ: Wiley-IEEE press, 2003, ch. 17, pp. 347–368.

[4] G. Karsai, J. Sztipanovits, A. Ledeczi, and T. Bapty, "Model-integrated development of embedded software," *Proc IEEE*, vol. 91, no. 1, pp. 145–164, Jan. 2003.

[5] P. Mosterman and G. Biswas, "A theory of discontinuities in physical system models," *J Franklin Institute*, vol. 335B, no. 3, pp. 401–439, 1998.

[6] G. Biswas, G. Simon, N. Mahadevan, S. Narasimhan, J. Ramirez, and G. Karsai, "A robust method for hybrid diagnosis of complex systems," in *Proc. 5th IFAC Symp Fault Detection Supervision Safety Technical Processes*, Washington, DC, June 2003, pp. 1125–1131.

[7] P. J. Mosterman and G. Biswas, "Diagnosis of continuous valued systems in transient operating regions," *IEEE Trans. Syst., Man Cybern. A*, vol. 29, no. 6, pp. 554–565, 1999.

[8] B. Duffield and A. Hanford, "Advanced life support requirements document," NASA-Lyndon B. Johnson Space Center, Houston, TX, Tech. Rep. JSC-38571, Rev. B, Sept. 2002.

[9] P. Bonasso, D. Kortenkamp, and C. Thronesbery, "Intelligent control of a water recovery system: Three years in the trenches," *AI Magazine*, pp. 19–43, 2003.

[10] K. D. Pickering, K. R. Wines, G. M. Pariani, L. A. Franks, J. Yeh, B. W. Finger, M. L. Campbell, C. E. Verostko, C. Carrier, J. C. Gandhi, and L. M. Vega, "Early results of an integrated water recovery system test," in *Proc 29th Int Conf Environmental Sys*, 2001.

[11] K. Brammer and G. Siffling, *Kalman-Bucy Filters*. Norwood MA: Artec House, 1989.

[12] G. Biswas, E.-J. Manders, J. Ramirez, N. Mahadevan, and S. Abdelwahed, "Online model-based diagnosis to support autonomous operation of an advanced life support system," *Habitation: An International Journal for Human Support Research*, 2004, to Appear.

[13] S. Abdelwahed, J. Wu, G. Biswas, J. Ramirez, and E.-J. Manders, "On-line hierarchical fault adaptive control for advanced life support systems," in *Proc 32nd Int Conf Environmental Sys*, 2004, to Appear.