WP4_006C.DOC

EUROPEAN COMMISSION
DG III INDUSTRY

INDUSTRIAL RTD PROJECT 20874

remafex

REMOTE MAINTENANCE for FACILITY EXPLOITATION

WP4 INTERNAL REPORT

SPECIFICATION OF EFI FUNCTIONS

Condition Monitoring of Hydraulic String and Turbine (S4)

BODY REFERENCE : WP4/006/REV.C

TITLE :	SPECIFICATION OF EFI FUNCTIONS - CONDITION MONITORING OF HYDRAULIC STRING AND TURBINE (S4)
SOURCE :	EFI
STATUS :	DRAFT
EDITOR :	ØYVIN SKARSTEIN
VERIFIED BY :	EIVIND SOLVANG
DISTRIBUTION :	WP4 TEAM MEMBERS
DATE :	22 SEPTEMBER 1997
NUMBER OF PAGES :	19 + APPENDIX
INTERNAL REFERENCE :	N:\DOK\13\REMAFEX\WP4\WP4_006C.DOC

CONTENTS

PREFACE

This document contains a description of the remote maintenance function «Condition Monitoring of Hydraulic String and Turbine». The specification is suitable for implementation in the C language, and the algorithms needed to implement the function are independent of other software packages.

The preparation of this specification presupposes a hydraulic string of radial structure, explained below.

1. OBJECTIVE OF FUNCTION

This function provides condition monitoring of the hydraulic string and turbine(s) of hydro power plants.

The condition in this context is defined as a set if hypotheses containing normal condition or one single «fault», i.e. either a process fault or a measurement device error.

Process faults are presented as a change in numerical value of one of the loss coefficients or the efficiency of the turbine. Measurement device errors are presented as a change of numerical value of the bias of a sensor (i.e. different from zero).

2. FUNCTION ENVIRONMENT

Based on measurements, the user program function performs state estimation, parameter estimation and hypothesis testing. Information from these three stages is used to present verbal and numerical information to the operator. The verbal information is a statement based on the hypothesis which most likely represents the actual situation in the hydraulic system.

Examples:

«The hydraulic string is normal within a tolerance of ... in loss coefficients.»

«The turbine(s) are normal within a tolerance of ... in efficiency.»

«The penstock is likely to have an increased loss of ... % relative to a normal value of ... s^2m^{-5}.»

...

«The pressure sensor just upstream of the main valve is most likely biased by ... m.»

etc.

In order to activate the function, a sequence of requests and responses are as follows.

Table 1 Sequence of request and responses necessary to utilize function.

Event	Request/Response (RQ/RS)	From	To
Perform monitoring	RQ	Maintenance operator (M.O.)	Control operator (C.O.)
Acknowledge	RS	C.O.	M.O.
Collect data - Specify number of samples and sample interval	RQ	C.O.	Data Acquisition system (D.A.S.)
Qualify data	RQ	C.O.	D.A.S.
Store data	RQ	C.O.	Data Base
Select set of hypotheses	RQ	M.O. (C.O.)	C.O.
For each hypothesis:			
- Perform state estimation	RQ	Function	Function
- Perform parameter estimation	RQ	Function	Function
- Calculate probability density function	RQ	Function	Function
- Select dominant hypothesis	RQ	Function	Function
Present results:			
- Verbal conversion of dominant hypothesis	RS	Function	C.O.
- State estimates (H, Q)	RS	Function	C.O.
- Parameter estimates (k_1, k_2, ...) (η of turbine(s))	RS	Function	C.O.
- Hypothesis probability	RS	Function	C.O.
Present conclusion (Essentially the same information as above)	RS	C.O.	M.O.

Internal events such as error messages, request for additional specification of task, i.e. limitation to subfunctions etc. are not included in the table above.

The results referred to in Table 1 are time series of recursive estimates and probabilities suitable for graphical time-plots.

3. INPUT DATA DEFINITION

3.1 USER INPUT AT REQUEST

1	DT	:	Sample interval	(s) (real scalar)
2	NSAMP	:	Number of samples	(integer scalar)
3	HYP	:	Selection of pre-defined hypotheses to be tested	(integer vector)

3.2 GENERIC CONFIGURATION OF HYDRAULIC STRING

The hydraulic string is assumed to have the structure shown in Figure 3.1.

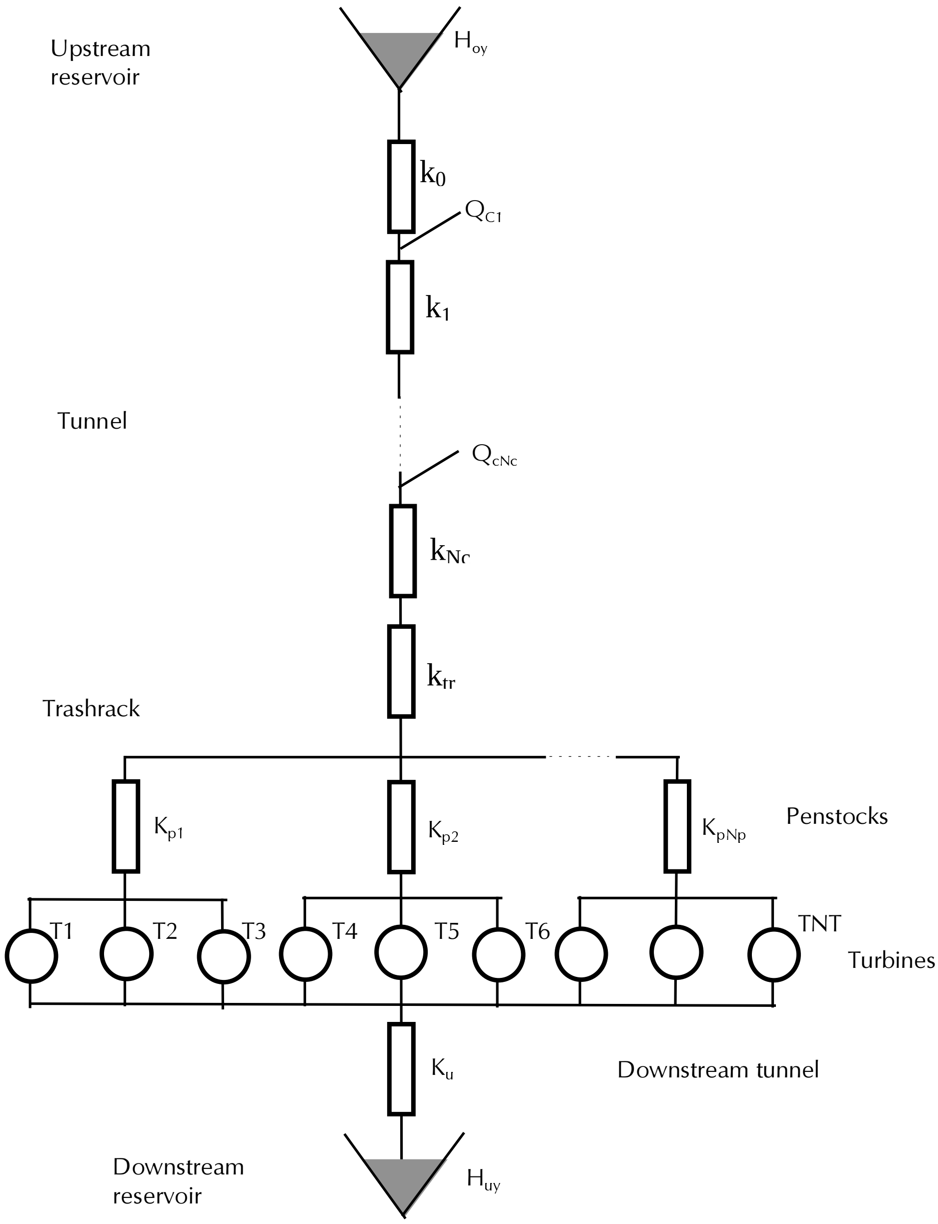

Figure 3.1 General radial structure of hydraulic string

The symbols used are:

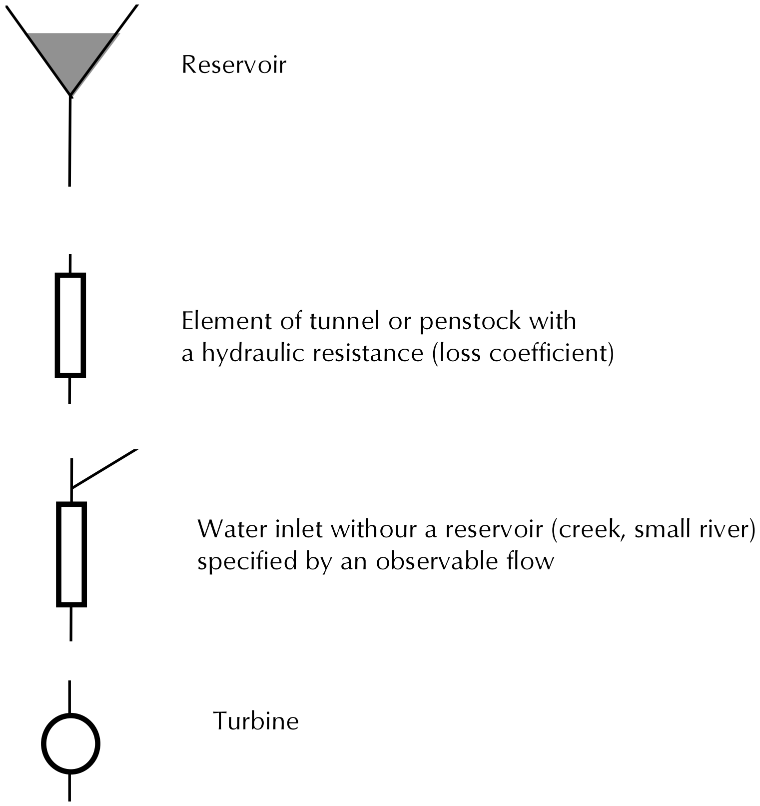

Figure 3.2 Symbols of hydraulic elements

3.3 PARAMETERS SUBJECT TO CONDITION MONITORING

Loss coefficients

Following the general structure of a single string multi machine radial system of Figure 3.1, the hydraulic losses are represented by the loss coefficients of the tunnels, penstocks and the thrashrack. The losses of head ΔH are described by

$$\Delta H_i = k_i Q_i^2 \text{ (m)}$$

this form being valid for all tunnel elements and singular losses as the thrashrack. If there are water inlets along the tunnel with significant flow, the tunnel losses must be subdivided into elements each described by an individual loss coefficient, because the losses are proportional to the square of the discharge. With no inlets, the tunnel can be described by a single loss coefficient.

The turbines

The method chosen for representation of turbine efficiency is the surface fitting using Chebyshev polynominals. The efficiency as a function of effective head H and discharge Q is then

$$\eta(Q,H) = F(XC,YC) = \sum_{i=0}^{K} \sum_{j=0}^{L} C(i,j) T_i(XC) T_j(YC)$$

Here η(Q,H) is the efficiency of operational point (Q,H) for the turbine. η is a number between zero and one. At bestpoint, η is typically 0,90 - 0,94 for Francis turbines. The function F is expressed in normalized coordinates XC and YC where

$$XC = \frac{2Q - (QMAX + QMIN)}{QMAX - QMIN}$$

$$YC = \frac{2H - (HMAX + HMIN)}{HMAX - HMIN}$$

This gives variation of XC and YC between -1 and +1 when Q and H vary between their minimum and maximum values.

The polynominals $T_i(x)$ are Chebyshev polynominals of 1.st kind, degree i and can easily be expressed in terms of recursion formula

$T_o(x) = 1$
$T_1(x) = x$
$T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x)$, i≥ 2.

The double sum is taken to the maximum degree K in Q and L in H which is necessary to represent the efficiency curves with the required accuracy. For most turbines K = L = 3 is sufficient and gives accuracy better than 10^{-3} in all operational points of the turbine.

The coefficient matrix C(i,j) is the actual data that need to be stored for each turbine. With K = L = 3, this is a 4x4 matrix. The convenience of this representation is that the polynominals are orthogonal, and if the nominal point of operation is chosen to be in the point (XC, YC) = (0,0) we see that the coefficient C(0,0) is the best-point efficiency of the turbine. All other coefficients represent the shape of the efficiency curve outside the best-point. For condition monitoring purposes, it is therefore possible to estimate only the C(0,0) parameter, assuming that a degradation of the turbine is not affecting the shape of the curve, only the level of it. This also means that the best-point efficiency will be observable regardless of whether the turbine runs at its best-point or not.

To determine the coefficient matrix C for each turbine, a set of efficiency measurements are needed, in which the whole range of operational points are covered. The actual algorithm to determine the best Chebyshev fit to those data is the EO2CAF routine of the Fortran NAG library.

Two other parameters are valuable for the observation of the state of the turbine. One is the Winther-Kennedy measurement. The other is the Torricelli law of the guide vane apparatus. The Winther-Kennedy measurement is a differential pressure sensor connected to the inner part and the outer part of the walls of the spiral case of the turbine. The flow creates a centrifulgal force on the wall as water runs in a circular trajectory through the case. The pressure difference measured is also a Q-square law

$$\Delta H_{WK} = k_{WK} \cdot Q^2$$

The coefficient of this equation is not subject to estimation (condition monitoring) but it is very valuable in order to establish a good estimate of the state vector, in which Q is the key variable.

The other measurement is the position of the main servo of the guide vanes. According to Toricelli's law, the discharge of the turbine is given by

$$Q = k_{T0} \cdot S_t \cdot \sqrt{2gH}$$

Where k_{T0} is Torrecelli's constant for the opening, S_t is the guide vane opening (m^2), g is the acceleration of gravity and H is the effective head across the turbine. This constant might be subject to estimation, to detect wear or damage of the guide vane apparatus.

The fundamental steady-state equation linking Q and H to the active power of the generator, which is an important measurement, is

$$P = \rho\, g\, \eta\, (Q,H) \cdot Q \cdot H \quad \text{(W)}$$

where ρ is density of water, g is acceleration of gravity, η is the efficiency of turbine and generator. The efficiency of the generator is a simple function of active power and can be set constant for most of the operational interval of the generator (close to 98 %).

3.4 TOPOLOGY OF ACTUAL PLANT

Following the general structure, the modelling of the hydraulic losses in the actual plant is defined by the topology parameters

Nc number of water inlets along the tunnel
Np number of penstocks
NT number of turbines. This number must be specified as a set of numbers, saying how many turbines associated to each penstock
$N_T = N_{Tp1} + N_{tp2} + \ldots\ N_{TpNp}$

Given these topology parameters, it is possible to write up the equations describing the conversion of energy and the losses in the plant.

Upstream tunnel

This is defined as running from the main inlet of the upstream reservoir, to and including the trashrack.

$$H_{oy} - H_p = \Delta H_{tr} + \sum_{i=0}^{Nc} \Delta H_i$$

$$\Delta H_i = k_i Q_i^2 \qquad \Delta H_{tr} = k_{tr} Q_{tr}^2$$

$$Q_{total} = Q_{tr} = Q_0 + \sum_{i=1}^{Nc} Q_i$$

Q_0 is discharge from reservoir, Q_i is discharge through inlet no. i, Hp is head on top of penstock(s). The discharge through the trashrack Q_{tr} is equal to the total discharge Q_{total} of the plant in this configuration.

Penstocks and turbines

The penstocks are assumed to be parallel structures, so we have

$$H_p - H_u = k_{p1}Q^2_{p1} + H_{Tp1}$$

$$= k_{p2}Q^2_{p2} + H_{Tp2}$$

$$\bullet$$

$$\bullet$$

$$= k_{pNp}Q^2_{pNp} + H_{TpNp}$$

Here, the parameters k_{p1}, k_{p2}, ..., k_{pNp} are the loss coefficients of each penstock, the loss being proportional to the square of the discharge, H_u is the head downstream of the common outlet of turbines, and H_{Tp1}, H_{tp2}, ..., H_{TpNp} are the effective head of each set of turbines associated with the respective penstock p_1, p_2, ... These effective heads enter into the state vector of the plant, together with each of the discharges through turbines.

Downstream tunnel

This is defined as running from the joint point of the draft tubes of turbines to the outlet of the downstream reservoir. This means

$$H_u - H_{uy} = k_uQ^2_{total}$$

It is assumed that there are no singular losses associated with each turbine outlet other than the losses in the turbine itself, and these losses are included in the efficiency function of turbines explained above.

H_u is the head just downstream of the turbines, and H_{uy} is the head (water level) of the downstream reservoir.

In case there is no downstream tunnel, we have $H_u = H_{uy}$ and there can be no extra hydraulic loss downstream of the turbine(s).

4. OUTPUT DATA DEFINITION

4.1 RANKED LIST OF HYPOTHESES

Ranked list of most likely hypotheses (integer vector) and the correspondent probabilities for a hypothesis to be the one to represent the actual condition in the plant. In most cases, by a single fault, or by a normal state, the top hypothesis should dominate (i.e. a probability close to one).

4.2 STATE VECTOR

The state vector of the hydraulic operation point is defines by the key variables of the turbines, Q_i ; i=1, ..., NT, where NT is the total number of turbines, and the difference in static head, H_{oy}-H_{uy}, between the upstream and downstream reservoirs.

The following structure is defined for the state vector $\underline{x}$:

$$\underline{x} = (Q_{T11}, Q_{T12}, \dots, Q_{T1NTp1}, Q_{T21}, Q_{T22}, \dots, Q_{T2NTp2}, \dots, Q_{TNTpNp}, H_d)^T$$

where $(\ )^T$ means the transpose of the row vector, and $H_d = H_{oy} - H_{uy}$ is the static head difference.

Example

With two penstocks and 3 turbines such that turbines 1 and 2 are connected to the first penstock and turbine 3 is under the second penstock, we get:

$$NT_{p1} = 2$$
$$NT_{p2} = 1$$

and the state vector:

$$\underline{x} = (Q_{T11}, Q_{T12}, Q_{T21}, H_d)$$

Given this state vector and the radial structure of the plant defined in section 3.2, and all parameters describing the hydraulic losses, then all other variables can be calculated from the state vector.

The effective head H_{Tp1} is common for turbine 1 and 2 and given by:

$$H_{Tp1} = H_d \text{ (upstream losses)}_{Tp1} - \text{(downstream losses)}$$

The upstream losses, referred to turbines on penstock 1 are:

$$\sum_{i=0}^{Nc} \Delta H_i + \Delta H_{tr} + k_{p1} Q_{p1}^2$$

In this expression, only the last term changes from penstock to penstock.

The downstream losses are:

$$k_u Q_{tot}^2$$

All symbols are explained in section 3.4, above.

This is the state vector that enters into the calculation of the normal state, i.e. no parameter estimation. This is the ordinary Kalman Filter calculation.

4.3 PARAMETER VECTOR

The parameters associated with thte condition monitoring are the loss coefficients of the hydraulic system: k_0, k_1, k_2, ..., k_{NC} of the tunnel (or just $k_t = k_0$ if there are no water inlets along the tunnel). Furthermore, the loss koefficient of the trashrack k_{tr}, the loss coefficients of the penstocks k_{p1}, k_{p2}, ..., k_{pNp} are subject to estimation, as well as the loss coefficient k_u of the downstream (tailrace) tunnel if there is such a tunnel. Then for each turbine, there is the coefficient C(0,0) defining the best-point efficiency, and then optionally the Torricelli constant of the guide vane apparatus opening k_{T0} might be an option for condition monitoring. The Winther-Kennedy constant of the spiral case can be taken as a fixed constant.

The parameters mentioned above are the process parameters. In addition, the monitoring method allows for estimation of faulty measurements, or bias in the sensors. There is one bias parameter for each sensor, and the measurement model for these bias errors is

$$y_m = y_r + b$$

where y_m is the indicated value, y_r is the true value and b is the sensor bias which is a systematic offset of the instrument.

5. DYNAMIC BEHAVIOUR

As mentioned in Section 3, the conditions of execution, operational modes etc., must be discussed carefully. The most important issue is the preselection of hypotheses to be subject to test.

The control operator or maintenance operator might have some a-priori information which could be helpful in establishing the list of hypotheses. In general, the number of hypotheses should be kept small, because of computing effort.

The operational mode of this function would be an intermittent one. Say, for instance, that the maintenance operator is interested in a condition monitor report every month or so. The function itself uses time series input data from the sensors. The successful estimation of parameters would need, say, some 500 samples from each sensor.

With sample interval of 10 seconds, the data acquisition function is active for 1 hour 23 min. The sample interval is not critical, since we are looking at a steady state process.

The example above leads to the following sequence in time.

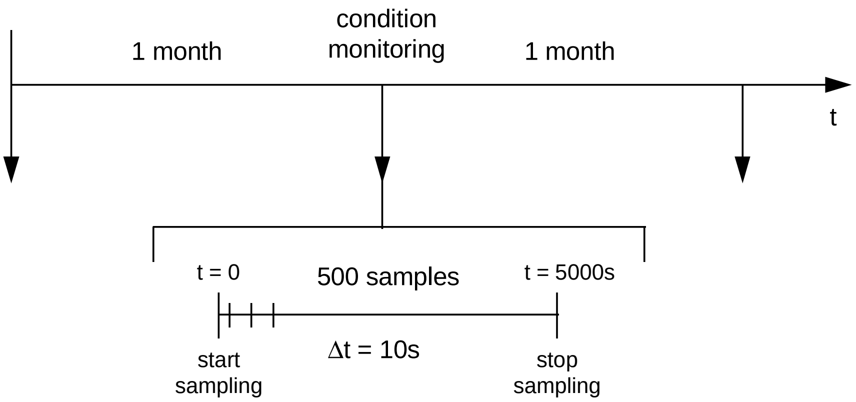

Figure 5.1 Time sequence of function (example).

During the interval (~ 1 month) between each time the function is activated, it is hibernating. Both the sample interval, number of samples and the hibernating period should be adjustable.

There is no need for automatic periodical activation of the function. It should be adequate for the maintenance operator/control operator to have the opportunity to manually request a run of the function.

6. DATA PROCESSING (ALGORITHMS)

Overall structure:

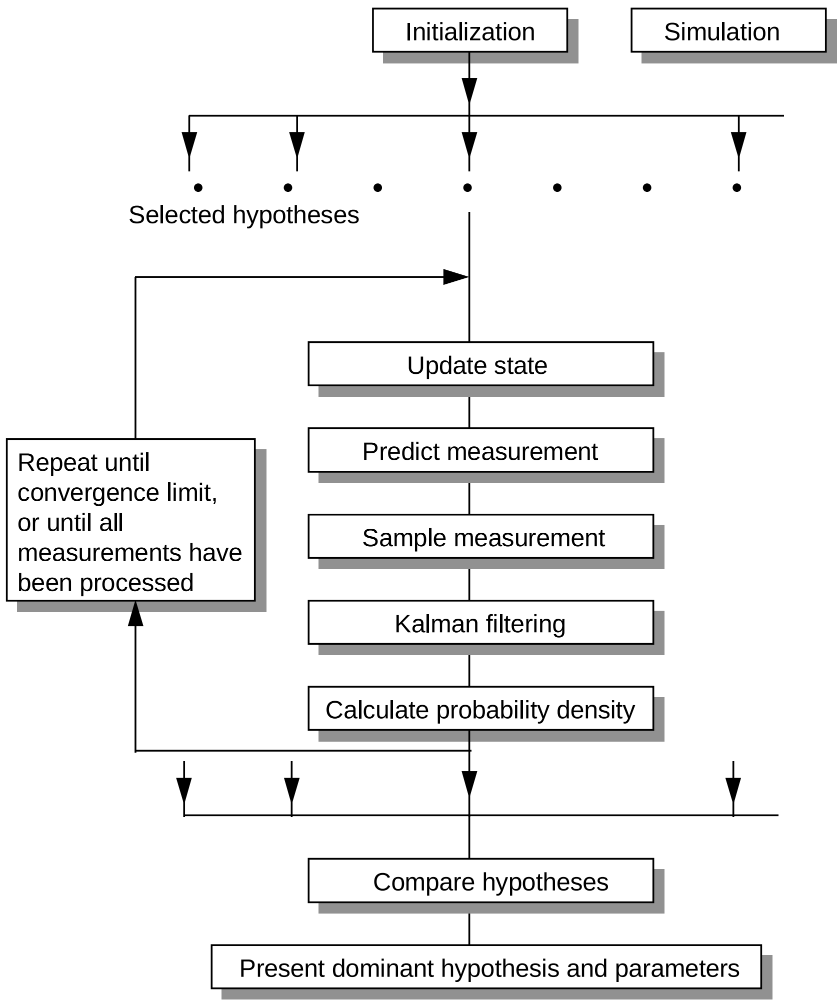

Figure 6.1 Overall program structure.

The sampling of measurements can be done in to ways:

1) A batch of say 500 samples is collected and stored on a datafile for off-line processing.

2) The sampling is made as a part of an on-line mode of the function. All predictions, filtering and calculation of probability densities are made in a recursive manner between each measurement sample.

The advantage of mode 2) is that the sampling can be stopped as soon as convergence is obtained.

All calculations are done in parallel, i.e. for each hypothesis. The last calculation is normalising and comparison of probability densities.

Simulator
The simulation «box» is intended for test purposes. All the algorithms should be tested off-line using synthetic measurements. This can be done using the simulation mode. This requires a detailed mathematical model of the hydraulic string and turbines of the plant. This model can be exactly the same as the measurement model of the state estimator.

Process failures can be simulated simply by altering the numerical value for some coefficient or efficiency. Sensor bias can be simulated simply by adding a constant value to one of the synthetic time series representing the sensor data.

It is strongly recommended to add (white) noise to the synthetic measurements, in order to simulate measurement noise and to make the estimation algorithms work properly.

The flow of data from the simulator and into the estimator of the condition monitoring system is illustrated by Figure 6.2 below

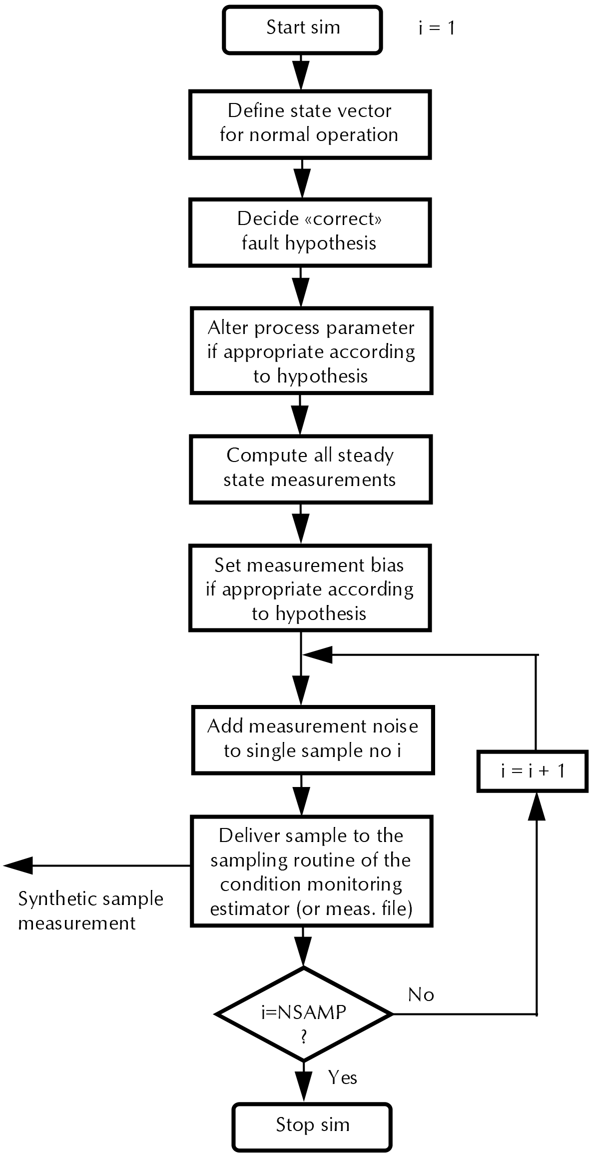

Figure 6.2 Flow chart for the simulator

The detailed description of each of the elements of Figure 6.2 is given in Appendix I, Algorithms.

Estimator
In the following a brief description of the elements is given.

- Initialization

  All matrices and vectors are initialised: State vector, measurement vector, parameter vector and covariance matrices.

  The state vector is extended with one additional element. This is a parameter subject to estimation. All other parameters are locked to their normal numerical values. The estimated parameter is uniquely defined by the actual hypothesis selected. One parameter for each hypothesis.

- Selection of hypothesis

  A hypothesis is defined such that a deviation from the normal (initial) state has occured. The normal state is defined as the set of process parameters being unchanged from their original values, and all the instruments being unbiased. The state vector can take on any value in the allowed range for the state vector elements. This means that for every turbine Q and H are between their minimum and maximum values if the turbine is running, or the Q is zero for each turbine at standstill.

  The hypotesis is simply a selection of those process parameters and/or instrument bias parameters which are suspected having changed their numerical value relative to the normal values (or the original values). The normal values can be opdated every time there has been a confirmed estimation of a change. The long-term trend can be recorded if these changes are followed over a period of time.

  The method of estimation of parameters used in this application is the Augmented Kalman Filter. This means that the state vector $\underline{x}$ is augmented by the parameter(s) which is (are) subject to suspicion of change in numerical value. The new state vector $\underline{x}$ is then the old vector $\underline{x}$ augmented by the number of parameters under suspicion (their value should be estimated based on the measurements):

  $$\underline{x} = (\underline{x}, \underline{p}_e)^T$$

  where $\underline{p}_e$ is a vector of parameters to be estimated. The vector $\underline{x}$ defines the hypothesis. The parameter $\underline{p}_e$ vector could in principle be all the process parameters and all the instrument biases, thereby making one single hypothesis carry all parameters. This would give a large state vector, large covariance matrises, and the observability of each parameter would be extremely low. It is therefore recommended to keep the number of selectable hypotheses low by allowing only single fault hypotheses to be selected. This means that the state vector would be augmented by a scalar only, namely the parameter under suspicion. This means that the maximum number of hypotheses is the number of parameters that could be estimated, which again is the number of process parameters (loss coefficients, efficiency coefficients C(0,0) etc.) plus the number of sensors.

  The restriction to single fault hypothesis only is a matter of choice. The reason for this recommendation is that faults in instruments occur independently and process changes occur either very slowly, increased wear, gradually obstructed ducts etc, or as a sudden change (large debris particles obstructing flow). By requesting the monitoring function with reasonably short time intervals relative to the expected changes in parameters, the assumption of single faults is a fair one. If this is regarded as a too strong restriction on the parameter estimation, a good alternative is to allow for double fault hypotheses. The $\underline{p}_e$ vector would then hold two elements, and the number of selected hypotheses would be:

$$\binom{\text{number of parameters}}{2}$$

Example

Say there are 12 process parameters and 17 sensors each prone to bias failure. The maximum number of hypotheses for single faults would be 12 + 17 = 29 in addition to the normal hypothesis. In case of double fault hypotheses, the maximum number would be:

$$\binom{29}{2} = 406$$

There is always a trade-off between flexibility and complexity. Flexibility in number of estimated parameters is achieved at the expense of computing time and observability.

- Update state

  Since we assume stationary conditions, the state is set equal to the previous state at every time step. This means that H and Q for each turbine and the parameter to be estimated are the elements of this steady state vector. They are updated by the Kalman filter.

- Predict measurements

  Given the state vector, the measurements prediction can be calculated as a known (non-linear) vector function of the state vector. A function for calculation of the efficiency of the turbine is used. The elements of the measurement matrix, a linearized version of the measurement function, is calculated.

- Sample measurements

  A new sample of measurements from each of the sensors are read from the batch file (mode 1), see above) or is collected on-line through the data acquisition system (mode 2), see above).

- Kalman filtering

  Performs a weighted updating of the state vector depending of the prediction error vector (predicted measurement minus actual measurement vector).

  The necessary calculations for the prediction of next state are the following:

  - apriori covariance of states (matrix)
  - Kalman filter gain matrix
  - aposteriori covariance matrix of states

- aposteriori estimate of state

- Calculate probability density

  The probability density of the numerical value that the measurement vector is calculated. This is the probability of a deviation between estimated values and values calculated from a noiseless model, referred to the model corresponding to the actual hypothesis.

- Comparison of hypotheses

  By the use of Baye's rule, the probability of hypothesis no. i is actually representing the present condition of the plant, is calculated. Here, it is assumed that the individual probabilities sum up to one, which means that one of the hypotheses is likely to be the correct one. This, obviously, depends on the skills and apriori knowledge of the operator. One should include the «Normal condition» hypothesis, to ensure that the normalisation is valid if no errors occur.

- Present the most likely hypothesis

  A comparison between the hypotheses results in tables in which the probabilities have numerical values between 0 and 1. If one of the probabilities approaches 1, this is taken as a positive diagnosis of the corresponding hypothesis being true. It might be the normal hypothesis, or it may be that a (single) error is localised.

  The condition is obviously that the space spanned by the individual hypotheses is a complete set. If none of the hypotheses dominates, and the normal hypothesis competes with the alternatives, there is probably no error in the system. If it is difficult to identify a single error, the reason may be that the number of hypotheses initially selected, were incomplete, or that a number of hypotheses describe the same situation (i.e. normal condition).

7. INTERFACES

7.1. OPERATOR INTERFACE

The program system is a stand-alone module suitable for coding in the C language. The reason that the C language is recommended, is that some of the submodules have already been coded, but not integrated in a complete system. Therefore, we believe that there might be some time and effort saved by choosing C.

All data into and out of the module are so far assumed to be held on flat files, so that the integration into an existing platform (Castor) should be as easy as possible.

In simulation mode, the user must run the program to produce synthetic data to test the performance of the estimator. In this way the user can inject errors in process or sensors. Prior to estimation, the user must select the hypotheses to be included in the comparison.

This is done interactively in a dialogue with the program. This dialogue can be implemented by clicking at alternatives in a comprehensive table of statements describing the individual hypotheses and their corresponding state model.

The results of the estimation are mainly time series of estimated states, parameters and probabilities. These data are suitable for graphical presentation as time plots.

7.2. SYSTEM INTERFACE

At this stage of development, the system interface is open. It must be specified in co-operation with those responsible for the overlying system (Castor).

8. ERROR MANAGEMENT

A general rule is that all input data should be subject to prequalification (limit check, plausibility or operator confirmation) before they are allowed to enter into any processing. The method itself is quite robust, provided that there are no anomalies in measurement data.

The calculation of probability densities requires adjustment of some convergence sensitive internal parameters. These are dependent on measurement noise, and must be tuned according to the requirements of the actual installation.

All error messages should be held on a separate file, and each individual message identified by a globally unique number which is referred to as a parameter by all the relevant functions and procedures. In this way it is easy to translate the dialogue into any language suitable for the operators. This principle also applies to the interactive dialogue controlling the running of the program.

A detailed list of error messages must be developed for each program module and submodule. This remains to be done, and it will be subject to further specification in the Appendix, following the software algorithm specification.

9. CONSTRAINTS

As mentioned in previous sections, there are no serious real-time constraints that must be fulfilled for this function to execute properly.

As long as the underlying assumption of steady state processes is valid, there is practically unlimited freedom of choice of sample interval length.

However, if the process contains dynamic modes, i.e., hydraulic (slow) oscillations, the process model (state space) must be modified accordingly, and additional parameters must be introduced which can describe the dynamical modes (oscillation frequency etc.). This requires no structural change in the algorithms, only that the state vector be expanded and the process model expressed as difference equations instead of steady state.

10. HARDWARE AND SOFTWARE REQUIREMENTS

The function can be implemented using standard PC hardware, a standard commercial data acquisition system with PC interface and a C compiler. No external software packages are required for the implementation of Kalman filter and hypothesis test algorithms.

11. TEST PLAN

It is strongly recommended that the simulator be implemented with the system. In this way it will be possible to test the system in a modular way, and the test, most importantly, can be made independently of data acquisition hardware.

The data acquisition system needs off-line testing before anything is installed in the pilot power plants.

If existing data acquisition is to be used, which is already installed in the control centre SCADA, or a combination of existing and new data collection hardware and software is to be used, then the test procedure must be specified and tailored to the actual installation. This point therefore has to be decided upon at a later occasion.

APPENDIX

Appendix 1: Detailed specification of algorithms and software sub-modules. This remains still to be completed. See the reference article below. Most of the basic algorithms are described herein.
Appendix2: Data tables

REFERENCE (attached document in appendix 3)

Skarstein, Ø.:
Condition monitoring of power plants using extended Kalman filtering.
Modelling, Identification and Control, 1988, vol. 9, no. 3, 149-163.

APPENDIX 1

«ALGORITHMS»

A1. ALGORITHMS

In the following all the modules of Figure 2, «Overall program structure», will be given a detailed description.

A1.1. INITIALIZATION

This module provides initialization of the model. This also includes the establishment by the user of the selected hypotheses subject to testing against measurements.

Hypotheses

The user is supposed to specify a limited number of hypothese given by the selection of parameters to be estimated. If the number of parameters to be estimated is zero, then the model is set to pure state estimation and a normal state is then assumed. This is a useful mode of operation if there is no evidence of a faulty condition, or if the goal is simply to estimate effective head and discharge for each turbine.

The user can choose between zero and five parameters to be estimated. They are chosen from a complete set of parameters covering the whole hydraulic system. These are the parameters listed in section 4.3 and 4.4. Section 4.3 covers process parameters (loss coefficients etc.) and section 4.4 covers the measurement error (bias) parameters. Note that the number of parameters depends on the number of penstocks and the number of turbines.

The dimension of the state vector, which in the normal state is twice the number of turbines (effective head and discharge for each turbine), is increased by one for each additional parameter to be estimated. It is recommended to keep the number of estimated parameters as low as possible, in order to maximize observability and to keep the processing time low. This can be accomplished by performing a condition monitoring often enough to ensure that the assumption of a single fault or a normal state is valid. When the dimension of the state vector increases from N to N+1, the dimension of all covariance matrices is increased to $(N+1)^2$.

Initialize filter

The Kalman Filter is initialized by specifying the initial state vector. This could in fact be a zero vector. It will be updated after a few time steps anyway.

Furthermore, the apriori covariance matrix of the states must be specified. This is generally a full matrix, but the initial value of it can be specified as a diagonal matrix in which each (diagonal) element is given a value which corresponds to the (assumed) variance of the corresponding element for the state vector. It depends on how well the elements are known in advance. For example, if the discharge of a specific turbine is known to be q within a margin of $q \pm \Delta q$, then the corresponding initial variance element of the apriori covariance matrix should be set to $(\Delta q)^2$. Default values should be suggested by the system. The apriori covariance matrix is denoted $\overline{X}(k)$, the initial value being $\overline{X}(0)$. Also, the initial value of the aposteriori covariance matrix $\hat{X}(0)$ is set equal to $\overline{X}(0)$.

Covariance matrix of process noise V

This matrix is a measure of the overall model error. It takes care of the errors we make by simplifying the description of the real hydraulic process into the model. The initial value V(0) of this matrix can be specified as a diagonal matrix. The elements should be given a value which reflects the assumed model error.

Covariance matrix of measurement noise W

This is simply a measure of the accuracy of the sensors. The matrix is initialized as a diagonal matrix, where the elements is the variance $(\Delta y)^2$ of each measurement.

Transition matrix Φ

Given the process model in discrete form

$$x(k+1) = f(x(k), p)$$

where p is the parameter vector, the transition matrix is defined as

$$\phi(k) = \frac{\partial f(x(k), p)}{\partial x^T(k)}$$

where x^T is the transpose of x. Φ(0) is the initial value of the transition matrix.

Measurement matrix

The measurement model

$$y(k) = g(x(k), p)$$

is linearized to give the measurement matrix D(k):

$$f(k) = \frac{\partial g(x(k), p)}{\partial x^T(k)}$$

where $x^T(k)$ is the transpose of x(k). Given the initial value of x, x(0), then D(0) can be calculated. This is necessary to start the Kalman Filter.

Kalman Filter gain

The Kalman Filter gain matrix K(0) can be calculated given all the initial values (for k = 0) or the updated values (for k > 0).

$$K(0) = \overline{X}(0)\, D(0)^T\, (D(0)\, \overline{X}(0)\, D(0)^T + W)^{-1}$$

A.1.2. UPDATE STATE

Based upon the predicted state (which was calculated from the previous updated state if the time step counter is greater that or equal to one. If the step counter is zero, this means the initial condition which then acts as the «updated» state at t = 0), the updated state can be calculated using the Kalman Filter gain matrix and the measurement prediction error (also called the innovation):

$$\hat{x}(k) = \bar{x}(k) + K(k)\left(y(k) - \bar{y}(k)\right)$$

where

$\hat{x}(k)$	is the updated (aposteriori) estimate the state vector (including estimated parameters).
$\bar{x}(k)$	is the predicted state vector, calculated at time k-1. If k=0, this is the initial state which must specified.
K(k)	is the Kalman Filter gain matrix.
y(k)	is the actual measurements at time k.
$\bar{y}(k)$	is the predicted measurement vector at time k, calculated based on information at time k-1. For k=0 this vector can be a zero vector, which means that the first measurement is taken as a large prediction error.

<u>Predict state</u>

All information up to time k is now available, and the prediction of the next state vector can be done as

$$\bar{x}(k+1) = \phi(k)\, x(k)$$

When the state $\bar{x}(k+1)$ has been calculated, the algorithm is made recursive by setting

$$k + 1 \rightarrow k$$

which means that all information up to time k has been used. The predicted state vector is then the key to all calculations at the next time step.

A.1.3. PREDICT MEASUREMENT

With given parameters and states up to time k-1, and the predicted state at time k, the measurement model is used to predict the measurement vector at time k, before the actual measurements are taken.

$$\bar{y}(k) = g\left(\bar{x}(k), p\right)$$

With this, the measurement matrix D can be updated by using the linearization from A1.1.

$$D(k) = \frac{\partial g}{\partial x^T}$$

The results of the estimation are now stored in a file. This file will contain states $X_1, \dots X_n$, the measurements $y_1, \dots y_m$, and the prediction errors in measurements $\varepsilon_1, \dots \varepsilon_m$, $\varepsilon = y - \bar{y}$, and n m is the number of state variables and measurement variables respectively.

A.1.4. APRIORI COVARIANCE $\overline{X}$

Calculates apriori covariance matrix for the states

$$\overline{X}(k) = \phi(k-1)\,\hat{X}(k-1)\,\phi^T(k-1) + V$$

A.1.5. KALMAN FILTER GAIN

Given the apriori covariance, the gain matrix is

$$K(k) = \overline{X}(k)\,D^T(k)\left(D(k)\,\overline{X}(k)\,D^T(k) + W\right)^{-1}$$

in which $\overline{X}$ and D are as defined above. W is the covariance matrix for measurement noise. Note that the Kalman gain increases with increasing V (through $\overline{X}$) and decreases with increasing W.

A.1.6. APOSTERIORI COVARIANCE

The aposteriori covariance is calculated as

$$\hat{X}(k) = \left(I - K(k)\,D(k)\right)\overline{X}(k)\left(I - K(k)\,D(k)\right)^T + K(k)\,W(k)\,K^T(k)$$

where I is the nxn identity matrix.

A.1.7. SAMPLE MEASUREMENT

If measurements are stored on file, this is simply a function that addresses the measurement vector at time k.

If measurements are sampled on-line from a data acquisition system, this module should activate the sampling. Note that it is important to keep track of real time and a constant sampling interval. This is only to maintain the possibility of adding dynamics to the process equations. In the present version, there is no dynamics involved, so the consequence of not having a constant sampling interval does not affect the estimates.

A.1.8. APOSTERIORI STATE ESTIMATE

Given the Kalman gain and the new measurement vector, the state vector can now be updated simply by

$$\hat{x}(k) = \bar{x}(k) + K(k)\left(y(k) - \bar{y}(k)\right)$$

where $y(k) - \bar{y}(k) = \varepsilon(k)$ is the measurement prediction error. ε(k) must be kept on a file for the post-processing of probabilities of each hypothesis. The state vector $\hat{x}(k)$ is used to predict the state vector for the next interval k + 1.

A.1.9. PROBABILITY DENSITY

This function is based on the Multiple Model Hypothesis Probability Test method (MMHPT) which is explained in the attached article in Appendix 2. For each iteration, the probability density for the innovation sequence (ε(k)) is calculated. The result, Ψ_ι, is a measure of the conditional probability of the prediction error ε under the hypothesis H_i:

$$\psi_i\left(k+1|k\right) = \frac{1}{\sqrt{(2\Pi)^m\,|R|}}\exp\left(-\frac{1}{2}\varepsilon^T(k+1)\cdot R^{-1}(k+1)\,\varepsilon(k+1)\right)$$

where R is the covariance matrix for the prediction error ε:

$$R(k) = D(k)\,\overline{X}(k)\,D^T(k) + W$$

Note that this matrix (or it's inverse) is used to calculate Kalman Filter gain matrix described in A.1.5.

A.1.10. REPEAT UNTIL STOP CRITERION

This function tests for a stop criterion, which is fulfilled if one of the following two situations occur:

1) The deviation between the predicted measurements and the actual measurements is less than a given limit, this means that the filter (including parameter estimation) has reached convergence. The limit should be set just above the expected value of the prediction error variance $<\varepsilon^T \varepsilon>$ as given by the process noise V and measurement noise W.

2) The total number of samples (NSAMP) has been processed.

If not, proceed with steps A.1.2. through A.1.10. until 1) or 2) occurs. When hypothesis H_i has been investigated, this function tests if there are any other hypothesis left to process. If so the whole process is reinitialized using the appropriate procedures of A.1.1., and the loop A.1.2. through A.1.10. is re-entered.

This is then repeated until all the initial hypothesis H_i, i = 1,2 ... , HYP have been processed and the corresponding sequences $\{\varepsilon\}$ have been stored in a file.

A.1.11 HYPOTHESIS PROBABILITIES

By using Bayes' rule, all the pre-specified hypotheses are now compared. For each hypothesis, this results in a table containing the estimated probability of the hypothesis being true, given all measurements up to and including time k + 1:

$$\hat{q}_i(k+1) = P_r\{H_i(k+1)|Y(k+1)\} = \frac{\psi_i(k+1|k)\,\bar{q}_i(k+1)}{\sum_{j=1}^{HYP}\psi_j(k+1|k)\,\bar{q}_j(k+1)}$$

In this expression, the apriori probability $\bar{q}_i(k+1)$ is the probability of the hypothesis H_i being true at time k. This apriori probability is calculated as a static prediction with a lower bound.

$$\begin{aligned}\bar{q}_i(k+1) &= \hat{q}_i(k) \quad &&\text{if } \hat{q}_i(k) \geq q_{0i} \\ &= q_{0i} &&\text{if } \hat{q}_i(k) < q_{0i}\end{aligned}$$

A typical value for the lower bound q_{0i} is 10^{-4}. ψ_i are the probability densities calculated in the module described in A.1.9.

A.1.12 OTHER FUNCTIONS

In addition to the functions described so far, some other functions are needed. These are especially for initializing the model based on the information in the configuration file. Also, functions for opening files, writing to and reading from files, and for interpreting information held on files, functions for matrix manipulation and string manipulation are needed.

APPENDIX 2

«DATA TABLES»

Item	**Hydraulic string and turbine**
Domain	**Supervision/Monitoring**
User Need	**To request hydraulic string and turbine condition**
Function	**Monitor hydraulic string and turbine condition**

No.	System	Subsystem	Parameter	State	Type	Unit of measure	Range	Alarm value	Trip value	Remarks	Data label
1.	Hydraulic string	Upstream Reservoir	Reservoir no.		Integer		1, 2, ...	≤ 0		config.	_HY_UR_#NUMB
2.	Hydraulic string	Upstream Reservoir	Head measurement covariance		Real	m^2	> 0	< 0		user inp.	_HY_UR_#H_WW
3.	Hydraulic string	Upstream Reservoir	Head		Real	m	(min, max)	OOR		meas.	_HY_UR_#H_X
4.	Hydraulic string	Upstream Reservoir	Head process covariance		Real	m^2	> 0	< 0		user inp.	_HY_UR_#H_VV
5.	Hydraulic string	Upstream Reservoir	Head state covariance		Real	m^2	> 0	< 0		init. by user	_HY_UR_#H_XX
6.	Hydraulic string	Upstream Reservoir	Head bias		Real	m	(min, max)	OOR		tunable par	_HY_UR_#HB_X
7.	Hydraulic string	Upstream Reservoir	Head bias process covariance		Real	m^2	> 0	<0		user inp.	_HY_UR_#HB_VV
8.	Hydraulic string	Upstream Reservoir	Head bias state covariance		Real	m^2	> 0	<0		init. by user	_HY_UR_#HB_XX
9.	Hydraulic string	Downstream Reservoir	Reservoir no.		Integer		1, 2, ...	≤ 0		config.	_HY_DR_#NUMB
10.	Hydraulic string	Downstream Reservoir	Head measurement covariance		Real	m^2	> 0	< 0		user inp.	_HY_DR_#H_WW
11.	Hydraulic string	Downstream Reservoir	Head		Real	m	(min, max)	OOR		meas	_HY_DR_#H_X
12.	Hydraulic string	Downstream Reservoir	Head process covariance		Real	m^2	> 0	< 0		user inp.	_HY_DR_#H_VV
13.	Hydraulic string	Downstream Reservoir	Head state covariance		Real	m^2	> 0	< 0		init. by user	_HY_DR_#H_XX

OOR = Out Of Range

Item	**Hydraulic string and turbine**
Domain	**Supervision/Monitoring**
User Need	**To request hydraulic string and turbine condition**
Function	**Monitor hydraulic string and turbine condition**

No.	System	Subsystem	Parameter	State	Type	Unit of measure	Range	Alarm value	Trip value	Remarks	Data label
14.	Hydraulic string	Downstream Reservoir	Head bias		Real	m	(min, max)	OOR		tunable p.	_HY_DR_#HB_X
15.	Hydraulic string	Downstream Reservoir	Head bias process covariance		Real	m^2	> 0	< 0		user inp.	_HY_DR_#HB_VV
16.	Hydraulic string	Downstream Reservoir	Head state covariance		Real	m^2	> 0	< 0		init. by user	_HY_DR_#HB_XX
17.	Hydr. string	Tunnel	No. of tunnel sections		Integer		1,2,....	≤ 0		user input	_HY_TL_#SECTIONS
18.	Hydr. string	Tunnel	Number of penstocks		Integer		1,2,....	≤ 0		user input	_HY_TL_#PENSTOCK
19.	Hydraulic string	Penstock	No. of turb. connected to penstock no. n		Integer		1,2,....	≤ 0		user input	_HY_PEn_#TURBINES
20.	Hydr. string	Tunnel	Tunnel sect. no.		Integer		1, 2, ...	≤ 0		Tunnel no. n	_HY_TLn_#NUMB
21.	Hydr. string	Tunnel	Tunnel inlet no.		Integer		1, 2, ...	≤ 0			_HY_TLn_#INLET_NO
22.	Hydr. string	Tunnel	Inlet flow		Real		> 0				_HY_TLn_#QCN
23.	Hydr. string	Tunnel	Static head meas. covar.		Real	m^2	> 0	<0		user inp.	_HY_TLn_#H_WW
24.	Hydr. string	Tunnel	Loss coeff. state		Real	s^2 m^{-5}	> 0	<0		tunable p.	_HY_TLn_#HC_X
25.	Hydraulic string	Tunnel	Loss coeff. process covariance		Real	s^4 m^{-10}	> 0	< 0		user inp.	_HY_TLn_#HC_VV
26.	Hydr. string	Tunnel	Loss coeff. state covar.		Real	s^4 m^{-10}	> 0	<0		init. by user	_HY_TLn_#HC_XX
27.	Hydr. string	Tunnel	Static head meas. bias		Real	m	(min, max)	OOR		tunable p.	_HY_TLn_#HB_X
28.	Hydraulic string	Tunnel	Static head bias proc. covariance		Real	m^2	> 0	< 0		user inp.	_HY_TLn_#HB_VV
29.	Hydraulic string	Tunnel	Static head bias state covariance		Real	m^2	> 0	< 0		init. by user	_HY_TLn_#HB_XX

Item	**Hydraulic string and turbine**
Domain	**Supervision/Monitoring**
User Need	**To request hydraulic string and turbine condition**
Function	**Monitor hydraulic string and turbine condition**

No.	System, Subsystem	Component, Subcomp.	Parameter	State	Type	Unit of measure	Range	Alarm value	Trip value	Remarks	Data label
30.	Hydroelectric set, Turbine Francis	Spiral case	Spiral case no		Integer		1, 2, ...	≤ 0		topology	_HS_TF_SC_#NUMB
31.	Hydroelectric set, Turbine Francis	Spiral case	Connected to penstock no.		Integer		1, 2, ...	≤ 0		topology	_HS_TF_SC_#TN
32.	Hydroelectric set, Turbine Francis	Spiral case	Sp. case diff. head meas. covariance		Real	m^2	> 0	< 0		user inp.	_HS_TF_SC_#HD_WW
33.	Hydroelectric set, Turbine Francis	Spiral case	Discharge		Real	$m^3\ s^{-1}$	(min, max)	OOR		meas./int.	_HS_TF_SC_#Q_X
34.	Hydroelectric set, Turbine Francis	Spiral case	Discharge proc. covariance		Real	$m^6\ s^{-2}$	> 0	< 0		user inp.	_HS_TF_SC_#Q_VV
35.	Hydroelectric set, Turbine Francis	Spiral case	Discharge state covariance		Real	$m^6\ s^{-2}$	> 0	< 0		init. by user	_HS_TF_SC_#Q_XX
36.	Hydroelectric set, Turbine Francis	Spiral case	Case speed coeff. state		Real	$s^2\ m^{-5}$	> 0	< 0		init. by user	_HS_TF_SC_#CS_X
37.	Hydroelectric set, Turbine Francis	Spiral case	Case speed coeff. proc. covariance		Real	$s^4\ m^{-10}$	> 0	< 0		user inp.	_HS_TF_SC_#CS_VV
38.	Hydroelectric set, Turbine Francis	Spiral case	Case speed coeff. state covariance		Real	$s^4\ m^{-10}$	> 0	< 0		init. by user	_HS_TF_SC_#CS_XX
39.	Hydroelectric set, Turbine Francis	Spiral case	Case speed coeff. parameter		Real	$s^2\ m^{-5}$	> 0	< 0		tunable	_HS_TF_SC_#CS_P
40.	Hydroelectric set, Generator		Generator no.		Integer		1, 2, ...	≤ 0		topology	_HS_GE_#NUMB
41.	Hydroelectric set, Generator		From penstock no.		Real		1, 2, ...	≤ 0		topology	_HS_GE_#TN
42.	Hydroelectric set, Generator		Active power meas. covariance		Real	$(MW)^2$	(min, max)	OOR		user inp.	_HY_GE_#P_WW

Item	**Hydraulic string and turbine**

Domain	**Supervision/Monitoring**
User Need	**To request hydraulic string and turbine condition**
Function	**Monitor hydraulic string and turbine condition**

No.	System, Subsystem	Component, Subcomp.	Parameter	State	Type	Unit of measure	Range	Alarm value	Trip value	Remarks	Data label
43.	Hydroel. set, Generator		Active power state		Real	MW	(min, max)	OOR		meas.	_HY_GE_#P_X
44.	Hydroel. set, Generator		Active power proc. covariance		Real	$(MW)^2$	> 0	< 0		user inp.	_HY_GE_#P_VV
45.	Hydroel. set, Generator		Active power state covariance		Real	$(MW)^2$	> 0	< 0		init. by user	_HY_GE_#P_XX
46.	Hydroel. set, Generator		Active power scale factor		Real		> 0	< 0		fixed par.	_HY_GE_#P_SCF
47.	Hydroel. set, Turbine Fr.	Guide vane	Guide vane no.		Integer		1, 2, ...	≤ 0		topology	_HS_TF_GV_#NUMB
48.	Hydroel. set, Turbine Fr.	Guide vane	From penstock no.		Integer		1, 2, ...	≤ 0		topology	_HS_TF_GV_#TN
49.	Hydroel. set, Turbine Fr.	Guide vane	Guide vane opening meas. covariance		Real	m^4	> 0	< 0		user inp.	_HS_TF_GV_#WW
50.	Hydroel. set, Turbine Fr.	Guide vane	Vane opening state		Real	m^2	(min, max)	OOR		meas.	_HS_TF_GV_#X
51.	Hydroel. set, Turbine Fr.	Guide vane	Vane opening proc. covariance		Real	m^4	> 0	< 0		user inp.	_HS_TF_GV_#VV
52.	Hydroel. set, Turbine Fr.	Guide vane	Vane opening state covariance		Real	m^4	> 0	< 0		init. by user	_HS_TF_GV_#XX
53.	Hydroel. set, Turbine Fr.	Guide vane	Vane opening scale factor		Real		> 0	< 0		tunable p.	_HS_TF_GV_#SCF
54.	Hydroel. set, Turbine Fr.		Efficiency state		Real		(0, 1)	OOR		tunable p.	_HS_TF_#EF_X
55.	Hydroel. set, Turbine Fr.		Efficiency proc. covariance		Real		> 0	< 0		user inp.	_HS_TF_#EF_VV

Item	**Hydraulic string and turbine**
Domain	**Supervision/Monitoring**
User Need	**To request hydraulic string and turbine condition**
Function	**Monitor hydraulic string and turbine condition**

No.	System, Subsystem	Component, Subcomp.	Parameter	State	Type	Unit of measure	Range	Alarm value	Trip value	Remarks	Data label
56.	Hydroel. set, Turbine Fr.		Efficiency state covariance		Real		> 0	< 0		init. by user	_HS_TF_#EF_XX
57.	Hydroel. set, Turbine Fr.		Efficiency form factors		Real matrix					Dimension 4x4 init. by user	_HS_TF_#EF_EFF
58.	Hydroel. set, Turbine Fr.		Efficiency form fact. dim.		Integer		(3, 4)	OOR		Dimension 4x4 init. by user	_HS_TF_#EF_FFD
59.	Hydroel. set, Turbine Fr.		Minimum discharge		Real	$m^3 s^{-1}$	> 0	< 0		user inp.	_HS_TF_#QMIN
60.	Hydroel. set, Turbine Fr.		Maximum discharge		Real	$m^3 s^{-1}$	> 0	< 0		user inp.	_HS_TF_#QMAX
61.	Hydroel. set, Turbine Fr.		Minimum head		Real	m	> 0	< 0		user inp.	_HS_TF_#HMIN
62.	Hydroel. set, Turbine Fr.		Maximum head		Real	m	> 0	< 0		user inp.	_HS_TF_#HMAX
63.											
64.											
65.											
66.											
67.											

APPENDIX 3

Paper:
«Condition monitoring of power plants using extended Kalman filtering»