Remote diagnosis by analyzing heterogeneous data

Joris GUERRY, François CHANEL, Nicolas TARDIEU

EDF R&D CHATOU Lab, 6 quai Watier, Chatou, France
{firstname.NAME}@edf.fr

Abstract

Based on a modernization project of its information system, EDF's hydropower sector is now able to merge several heterogeneous databases to produce new applications. EDF Research and Development produced a new software to assess the impact of maintenance actions on performance of equipments. In practice, the tool allows, on the one hand, to display several time series of sensors between two dates, and on the other hand, to superimpose on the same graph the relevant maintenance acts, linked to the displayed sensors. We use the new efficient structuring of sensor nomenclatures and maintenance data to sort by relevance the maintenance events to be displayed first, based on expert knowledge.

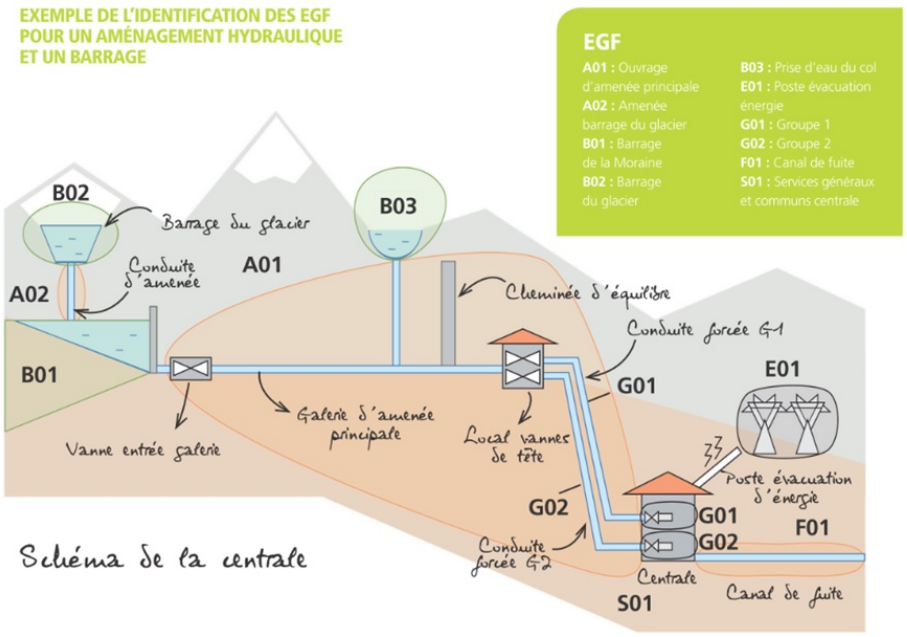

Figure 1: Schematic diagram of a hydraulic installation and name of the main elements (in French)

1 Technical context and need

1.1 Separation of process, maintenance and alert data

Historically, in the Production and Hydraulic Engineering Division of Electricité de France (EDF), process, maintenance and operation alert data were recorded and consulted in totally separate IT tools.

The data capitalized in these databases were therefore exploited separately, and no advantage was taken of any cross-checks and cross-references of information between databases, which could have led to a better understanding of events, a better anticipation of operational fortuitous events, and a better monitoring of the effect of the maintenance applied. One of the objectives of a data cross-referencing tool is to remove these information silos to allow enriched analyses by consulting heterogeneous databases.

1.2 Separation of operations, monitoring and maintenance units

As EDF's hydraulic division is large, it is segmented into different specialist units: some units specialize in the operation of hydraulic installations, others in the programming, implementation and operational monitoring of maintenance operations, and others in the e-monitoring of hydraulic production facilities.

While these different units cooperate effectively on a daily basis on many issues, each one carries a lot of information that is specific to its business lines, and the sharing of information between units cannot be perfect and complete on an ongoing basis. In particular, **some scheduled maintenance events may not be known to the e-monitoring units**, while these operations then generate monitored signals which, if not placed in the context of ongoing maintenance, may lead to a misinterpretation of abnormal situation and lead to a monitoring alert. These situations of imperfect interpretation then require additional communication time between operations, maintenance and monitoring engineers, whereas a centralizing digital tool could have made relevant information immediately accessible to everyone and avoided any incorrect interpretation of the signals read and operations in progress.

1.3 Standardization of equipment names (ECSH codes)

In recent years, EDF's hydraulic division has carried out a major project to standardize and homogenize the various tools of its information system, which has made it possible to standardize the model for forming the names of hydraulic power plant components (at least for high-power hydraulic sites). The template of names obtained is called ECSH codification (for EDF Coding System Hydraulic).

This standardization was an essential prerequisite for the project to create a heterogeneous data cross-referencing tool, because it was essential that an object present in the operating databases (as part of the time series that are measured and recorded via the various sensors with which it is equipped) as well as in the maintenance databases (as part of the various maintenance operations that have been carried out to ensure its reliability over time) bears the same name within these different tools: this makes it possible to identify which data are relevant to be processed together. This coincidence of names was therefore the first prism used to group heterogeneous data together.

Figure 1 shows the ECSH names used to designate the main structures of a hydropower plant : these names then serve as roots for the names of all the sub-components of each major structure. Figure 2 shows the decomposition of a tag in our hydraulic data historian. This tag is the name of a sensor time series, and it is directly derived from the ECSH code name of the component on which said sensor is placed. Figure 3 shows the decomposition of a component name associated to a maintenance event (dated text data) in our hydraulic Computerized Maintenance Management System (CMMS) named Gem'Hydro. As this is actually the same component in both cases (the component on which the sensor is placed and whose data are recorded, and the component on which a maintenance operation has been performed and then recorded), one can see that the parsing of the 2 names reveals the same construction by tree structure, which will later allow to group and compare these data.

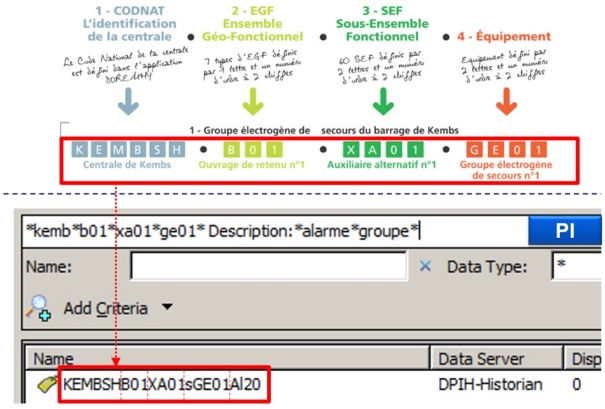

Figure 2: Parsing of the name of a PI tag linked to the equipment tree in which the monitored equipment is located

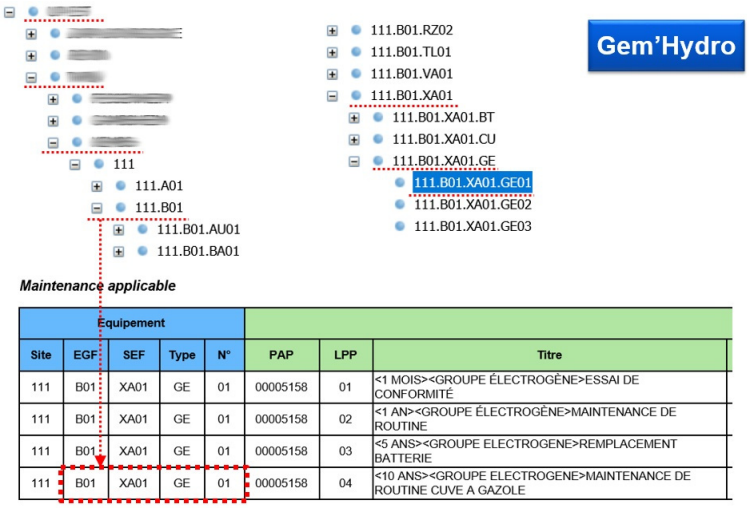

Figure 3: Parsing of the name of a maintenance event in Gem'Hydro related to the equipment tree in which the maintained equipment is located

1.4 Target audience and expected benefits

The first audience targeted by this tool is Engineers performing e-monitoring of installations in CReEX (regional e-Operating Centres) and Engineers of the operator support teams.

The expected benefits of the tool include:

- The centralization of heterogeneous information (here: time series and structured textual data) in a single tool, thus facilitating their cross-analysis. The searchable data will be enriched as decisions are made to release other databases into the tool.
- The possibility of making visual comparisons before / after maintenance, in terms of sensor signals evolutions, and thus to assess qualitatively and quantitatively the effects produced by maintenance operations
- The possibility of directly seeing the quantitative effects of operating recommendations
- **Allowing a better circulation and exploitation of relevant information for the monitoring of installations, which also limits the risk of false alarms, limiting unnecessary solicitations and exchanges between monitoring, operation and maintenance units by sharing the same context.**

2 A new cross-referencing module

This tool is currently being developed as a Proof of Concept, in order to demonstrate the interest for the operational divisions in developing a data cross-referencing tool that allows them to make full use of their very large data assets (time series, highly structured document base, etc.). As such, it is still subject to several restrictions: only data from a few hydraulic sites are dumped there during the current test phase, the current IT architecture is not yet optimized to support a large number of concurrent requests. Depending on feedback and economic evaluation (effectively saved engineer time and avoided costs), additional developments may be considered for a truly robust industrial deployment.

In order to reduce EDF workload, we merged our development into an existing visualisation tool from another EDF entity, SuperViz'Orte, named after the first site testing it. Our data cross-referencing software then took the form of an additional module added to this more general tool for consulting and analyzing hydraulic time series. This module is shown in Figure 6.

2.1 Time series of our sensors: PI

The database used to capitalize the time series of hydraulic production facilities is the PI OSIsoft data historian, a robust and widely proven industrial solution. The PI OSIsoft Application Programming Interface

(API) allows to extract the PI tags corresponding to a set of parameters, which then allows the user to precisely choose the time series to be displayed over the period of his choice. The native services of the PI suite then allow you to directly request either the raw time series (extracted at the real time step of measurement and recording), or the average values calculated on the fly (limiting the total volume of data transmitted, when the exploitation of the data at a time step as narrow as that of the raw data is of no interest).

2.2 Act of maintenance database : Gem'Hydro

The CMMS tool used by the hydraulic division is an Oracle database called Gem'Hydro. This database imposes a strong structure in the recording of the various work orders: this structure has therefore made it possible to develop fairly fine weights for the selection of relevant maintenance tasks to be proposed for display, in addition to the process time series graphs.

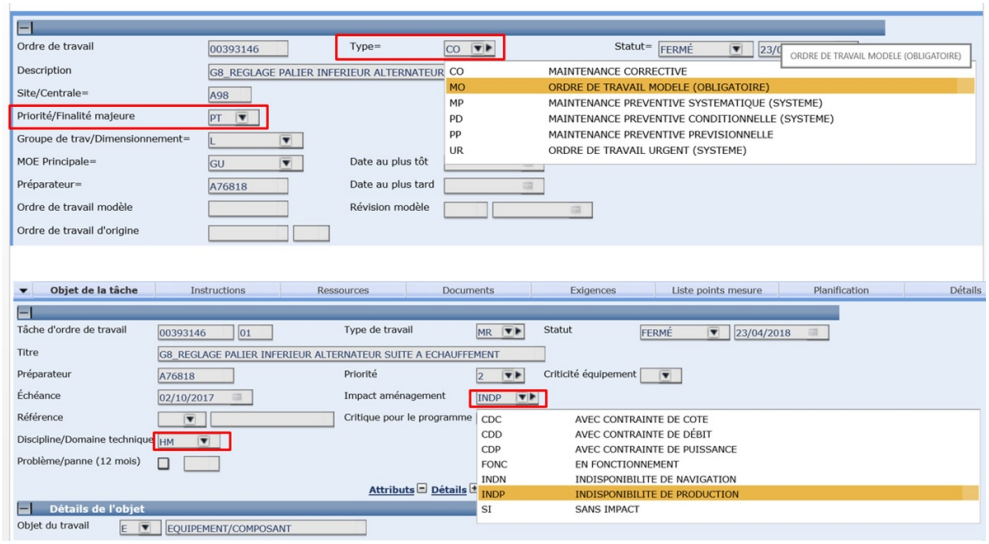

Figure 4: View of the different fields that apply to maintenance tasks registered in Gem'Hydro

Figure 4 shows a view of the different fields that must be filled in when recording a maintenance order or task. These different fields allow you to specify:

- The preventive or corrective nature of the maintenance action carried out
- The general purpose of the intervention (hydraulic safety, maintenance of the assets, increase in performance, etc.)
- The impact on production (impossible production, reduced river navigation, etc.)
- The general field of competence concerned (civil engineering, electrical engineering, control command, etc.)

In addition, the tool allows the recording of a comment in free entry. Operators sometimes comment extensively on the intervention performed and the salient information to be retained like shown on Figure 5. This input is to take into account with precaution as mistakes or abbreviations can be used.

The feedback from this database after several years of existence is double-edged: its robust design and the fairly exhaustive mandatory entry lead to a very good level of reliability of the information contained in most fields (only a few of the available fields had to be discarded due to a lack of satisfactory reliability). On the other hand, the relative "heaviness" of the associated interface and its not very ergonomic nature lead to a minimal input by the maintenance units: only strictly mandatory information (major work) is included, whereas ideally one would like to see the slightest small maintenance intervention entered completely.

Finally, since the system name tree structure through which maintenance tasks are recorded is the same as the one around which process PI tag names are constructed (both following the templates imposed by ECSH coding), it is relatively easy to associate the right maintenance acts with the right process time series curves.

Historique des activités								
Date de fin de réalisation	Objet du travail	Liste d'équipements	N° TOT	Type d'OT	Libellé de tâche	Commentaire de fin de tâche	Références documentaires du compte rendu	
01/06/2018	A98.G04.GU04.PB01		00365469-01	MP	<S><PIVOTERIE>CONTRÔLE PRESSION POMPE INJECTION_ME	ras		
01/06/2018	A98.G04.TU04.JS01		00450208-01	CO	REQUALIFICATION SUITE REMPLACEMENT JOINT A FUITES G4	Requalifs. ok => T°C et débit joint à fuite ok.	H-41554603-2018-000143	REQUALIF. G04 T°C + JOINT A FUITE_01062018
31/05/2018	A98.G04.AL04.AT04		00365448-01	MP	<R><ALT>VISITE X-MA-08-14 CTRL ROTOR ET INDEX_EL	Contrôle rotor réalisé, résistances de connexions, équi-répartition et index de polarisation sur le stator. R.A.S Présence d'un trou en partie baesse de la masse polaire sur le pôle 3 Suite au vernissage du rotor par OMEXOM ezn 2016 - aucune trasse d'extrusion de matière sur l'ensemble des pôles	H-41554603-2018-000141	20183005 G04 CONTRÔLE ROTOR
31/05/2018	A98.G04.AL04.AT04		00445255-01	MP	<ALTERNATEUR>MAINTENANCE LOURDE TÊTE D'EXCITATION_EL	Entretien lourd réalisé le 30/05/2018 suite à travaux EIM joint à fuite, contrôle rotor...: Aspirateur nettoyé, filtre remplacé, R.A.S. Isolement aprés remontage 1.5Gohm		
31/05/2018	A98.G04.TU04.TB04		00462280-01	CO	REMPLACEMENT DU JOINT À FUITE DU G4	Remplacement du joint fuite effectué	H-41554404-2018-000003	RT GMH - 2018 - G4 - REMPLACEMENT DU JOINT À FUITE
						Entretien léger à 500h + nettoyage des têtes de bobines pour essayer de faire remonter l'isolement.		

Figure 5: Comment fields associated with the maintenance tasks performed

2.3 Prioritization of maintenance operation reports

The tool is designed so that the user first chooses the production site and the sensor data (time series) he wants to display, as well as the start and end dates of the display. This forces him to specify which local group of hydraulic production units, site, production group and components he wants to display the sensor data of. The first step is to display the time series graphs, as shown in the top of Figure 6.

Once these first choices have been made, all the maintenance reports corresponding to this local grouping of units are displayed below the time series (see array in 6). However, the order in which they are proposed and displayed to the user is calculated according to a weighting that takes many criteria into account. **This weighting is intended to promote the immediate display of the most relevant reports, related to the theme explored by the user and the selected options.**

- Very high priority is given to reports on the same equipment as those whose sensor times series graphs are displayed, as well as those that are time-stamped between the start and end dates required for the display of sensor data
- The presence of filters (editable on the right of Figure 6) also makes it possible to modify the weighting of the reports: these filters include the attributes of the maintenance tasks in the Gem'Hydro database (presented in Figure 4), i.e. their selection will reinforce the priority of the reports presenting these same selected attributes. Several attributes have been checked by default (as per instructions from the operational units). It is also possible to add keyword search which would still prioritize acts of maintenance without rejecting any.

Then, the user is led to make a selection of reports relating the maintenance events that seem most relevant to the theme he is trying to explore (as long as the priority weighing of the reports is effective, and the selection criteria expressed are relevant, these are the reports that have been proposed to him first). The selected reports are then displayed as vertical bars in the time series graph window, as shown in Figure 6.

This cross-display therefore makes it possible to directly visualize whether the selected maintenance events correlate with the displayed time series. In the case studied in Figure 6, we displayed concomitantly:

- Time series of cold air temperatures of the generator in group 1
- Maintenance reports on the refrigeration system of the same alternator

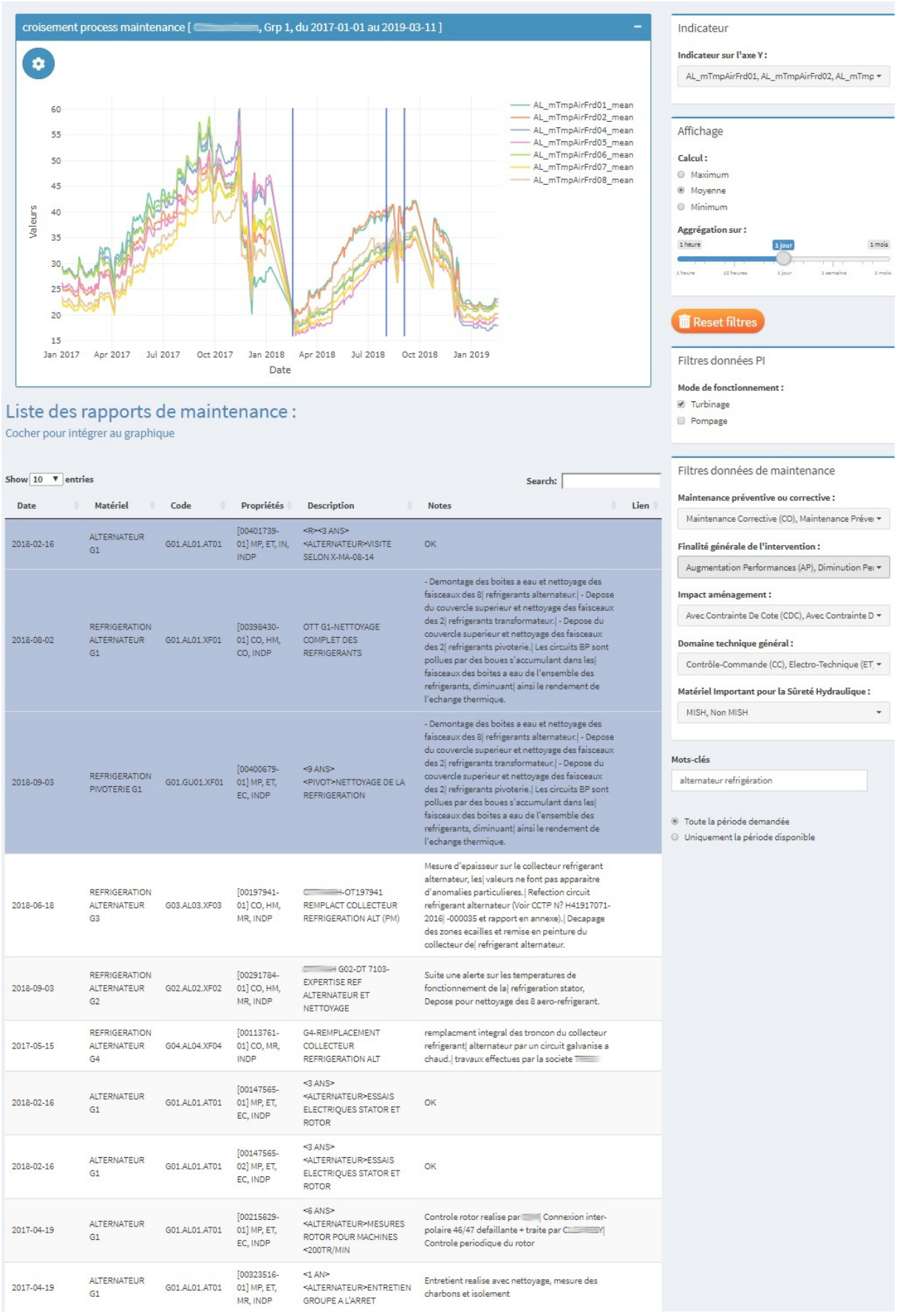

Figure 6: SuperViz'Orte cross-reference module screenshot (in French). The context restricts time series to alternator cold air temperatures and maintenance operations related to the alternator refrigeration system.

It can be seen directly on the Figure 7 that the maintenance operations carried out have the expected effect on the evolution of the component: at each maintenance operation on the refrigeration system, the alternator's cold air temperatures drop afterwards. Looking at the global Figure 6 might give the impression that the intervention causes an increase in cold air temperatures. In reality, the graphs display is partially misleading: since no operating data are recorded during the shutdown during which maintenance operations are carried out, straight lines are drawn between the actual data points, before and after maintenance, which may lead to the impression that there is a gradual decrease that stops as soon as the intervention on the refrigeration system is carried out, whereas in reality it is these operations that cause a drop in generator cold air temperatures, observed at the end of the intervention by measuring temperatures during the resumption of production (temperatures which then actually rise again and will lead to other acts of maintenance).

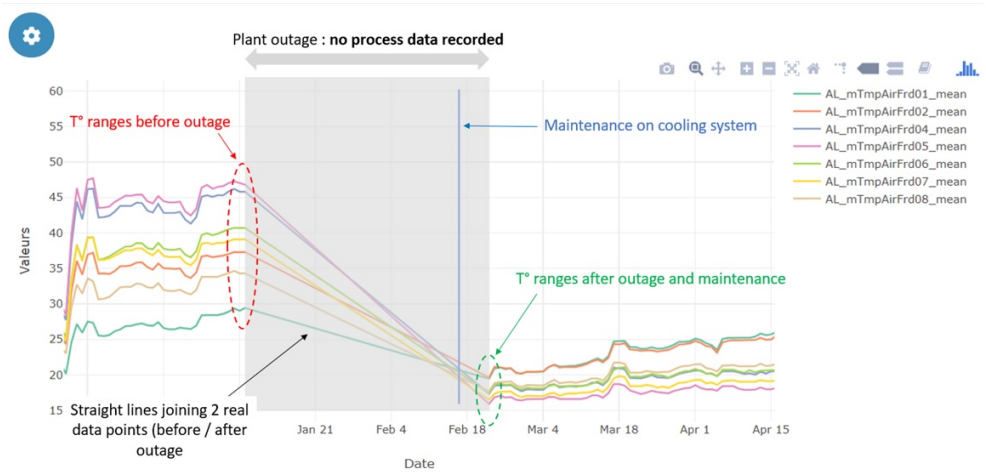

Figure 7: Focus on the first maintenance intervention: the plant outage and the maintenance performed are indeed at the origin of the temperature drops

3 Conclusion and perspectives

The tool has already been tested among the maintenance and monitoring units of EDF's hydraulic fleet. Feedback is considered positive: several relevant use cases (as shown in this article) have already been identified, showing that this application will be of real interest to these units in the future.

It is considered to add other different databases to the tool, so that it can probe a wider range of "events" to be displayed together with operational data: in particular, it is thought to add data on hazard and alert detections issued and capitalized by the operating and monitoring units. The expected benefits would be numerous: to note if the alerts are followed by actions (and to improve their consideration), to note if they have sufficiently anticipated the occurrence of possible operational hazards, ...

The weighing of relevance by which maintenance reports are prioritized is currently of relatively modest complexity. We could try to probe the documentary collection (in particular thanks to the free entry comment field) with more elaborate text mining tools, making it possible, for example, via the entry of a query expressed in French sentences and summarizing the general theme explored by the user, to propose a more robust order of relevance of the maintenance reports.

For the moment the interface only works in one direction: the user chooses which sensor data to display, then which maintenance reports to display among those proposed. One could consider an inverse operation, which would fully automate the selection of reports: according to the complete request expressed by the user, as well as the detection of salient events among the graphs which would then be exploited as ground truth (growth inversion, sudden changes in growth rate, etc.), propose to display the maintenance reports relevant for the context studied and time-stamped in the vicinity of these salient events of the operating data.

Acknowledgement

We thank our colleagues from EDF Hydro, and especially the CReEX UP Alpes who provided insight and expertise that greatly assisted the real heart of the developed module.

We thank Jérôme Boudon from EDF CIH for his assistance with the PI software part.

We would also like to show our gratitude to the "Usine Data Analytics pour les Producteurs" who gave us access to the Gem'Hydro database and provided us with their best R Shiny developers to merge our module into their visualization tool SuperViz'orte. A great thanks to Antoine De Caminel, Da Lan Nguyen, Laura Rouhier and Yann Renaud.

It was a real pleasure to work with all of them.