IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON COMMUNICATIONS IN HARSH ENVIRONMENTS

Received September 4, 2020, accepted September 19, 2020, date of publication September 24, 2020,
date of current version October 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3026348

Slow Degradation Fault Detection in a Harsh Environment

ANTHONY D. CHO[1], RODRIGO A. CARRASCO[1,2], (Member, IEEE), GONZALO A. RUZ[1,2,3], (Member, IEEE), AND JOSÉ LUIS ORTIZ[4]

[1]Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Santiago 7941169, Chile
[2]Data Observatory Foundation, Santiago 7941169, Chile
[3]Center of Applied Ecology and Sustainability (CAPES), Santiago 8331150, Chile
[4]Atacama Large Millimeter/Submillimeter Array (ALMA), Santiago 7630355, Chile

Corresponding author: Rodrigo A. Carrasco (rax@uai.cl)

This work was supported in part by the ANID FONDECYT Project under Grant 1180706.

ABSTRACT The ever increasing challenges posed by the science projects in astronomy have skyrocketed the complexity of the new generation telescopes. Due to the climate and sky requirements, these high-precision instruments are generally located in remote areas, suffering from the harsh environments around it. These modern telescopes not only produce massive amounts of scientific data, but they also generate an enormous amount of operational information. The Atacama Large Millimeter/submillimeter Array (ALMA) is one of these unique instruments, generating more than 50 Gb of operational data every day while functioning in conditions of extreme dryness and altitude. To maintain the array working under extreme conditions, the engineering teams must check over 130,000 monitoring points, combing through the massive datasets produced every day. To make this possible, predictive tools are needed to identify, hopefully beforehand, the occurrence of failures in all the different subsystems. This work presents a novel fault detection scheme for one of these subsystems, the Intermediate Frequency Processors (IFP). This subsystem is critical to process the information gathered by each antenna and communicate it, reliably, to the correlator for processing. Our approach is based on echo state networks, a configuration of artificial neural networks, used to learn and predict the signal patterns. These patterns are later compared to the actual signal, to identify failure modes. Additional preprocessing techniques were also added since the signal-to-noise ratio of the data used was very low. The proposed scheme was tested in over seven years of data from 132 IFPs at ALMA, showing an accuracy of over 70%. Furthermore, the detection was done several months earlier, on average, when compared to what human operators did. These results help the maintenance procedures, increasing reliability while reducing humans' exposure to the harsh environment where the antennas are. Although applied to a specific fault, this technique is broad enough to be applied to other types of faults and settings.

INDEX TERMS Echo state networks, predictive maintenance, condition monitoring, fault detection, harsh environments, observatories.

I. INTRODUCTION

In the last couple of decades, the complexity of ground telescopes has increased exponentially. A new generation of industrial-scale telescopes are being constructed, all of which share similar operational difficulties: multiplicity of instrumentation, high-levels of automation, multiple sensors for metrology, and remote management. All of these parts are continually generating massive amounts of operational data, which, although dwarfed in comparison to the scientific data produced, can still be in the order of hundreds of gigabytes every day. Due to the weather and sky requirements, most of these telescopes are located in remote locations and suffering harsh environmental conditions, making their complexity even more challenging. Severe dryness, extreme temperatures, and high-altitudes (and thus low oxygen levels) are just some of the conditions in which these instruments operate.

In order to ensure the high level of performance expected from these telescopes, constant monitoring of the different subsystems is required. Maintenance engineers generally

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang.

made this task in the older telescopes since the number of monitoring points was minimal. In contrast, the new generation of industrial-scale telescopes have hundreds of thousands of monitoring points. Hence, they generate operational data as never before, making previous monitoring efforts obsolete or utterly impractical due to how time-consuming it would be. To achieve the needed level of performance, automated data processing techniques are required in modern instruments, to comb through the gigabytes of data generated daily, and detect possible failure patterns, implementing predictive maintenance methodologies. These automated fault detection systems help maintain the expensive infrastructure and protect the engineers from exposure to the harsh environment, reducing the requirements of in-site revisions. Hence, sound predictive maintenance systems are essential to make the new scientific discoveries of these telescopes, feasible.

A. LITERATURE REVIEW

Fault detection systems have been developed since the early 1970s [1], [2], as an essential part of automatic control systems. In this work, we refer to fault detection as the process of determining if a system or subsystem has entered a faulty operation mode, i.e., a mode different from the normal operating conditions. This procedure is critical to ensure that things are running correctly. Fault detection tools are particularly useful in predictive maintenance systems to improve the use of expensive equipment.

Fault detection procedures can be divided into three main categories: signal processing techniques [3]–[6], model-based techniques [7]–[9], and data-driven ones [10]–[16]. In this work we will focus in the latter, where data will drive the identification of normal and faulty modes of operation. See [1] for a general description of fault detection and diagnosis systems.

Autonomous fault detection systems have been developed in many different areas. In the case of model-based techniques, authors have applied these techniques to physical systems that are relatively easy to model. For example, Korasgani *et al.* [7] develop a model-based fault detection scheme for a system of two tanks with two valves, with the capacities and resistances of tanks and valves as model parameters. Using their method, they can detect faults that appear in the valves. They also define a detectability ratio, a measure of residual detectability performance, providing a means to find the most sensitive and robust residual for the uncertainties at different regions of the system. They also test the effectiveness of this approach and show that it is possible to achieve better detection by using more than one residual. In an application to modern telescopes, Ortiz and Carrasco [8], [9] developed a framework of fault detection and diagnosis, using a bank of Kalman filters to detect a specific type of slow degradation faults. They tested their scheme with real fault data from ALMA, showing excellent accuracy. Their main drawback is that a model of the faulty and non-faulty systems is required, something that is time-consuming and hard to generalize. In this work, we tackle the same type of fault but using a data-driven approach that scales and generalizes much better.

One of the main techniques used in data-driven fault detection is to apply different architectures of Artificial Neural Networks (ANN). For example, Wootton *et al.* [13] developed a fault detection system for bridge structures. In their work, they study the structural health of a footbridge using data time series collected from temperature and tilt sensors located in strategic places. They show that, by using echo state networks to replicate the bridge's behavior, they can detect faults in the bridge's structure. To do so, they compare the tilt sensor measurements with their predictions from the temperature data. In another application using echo state networks, Morando *et al.* [14] introduced an approximation to fault diagnosis of fuel cells stacks, using a variation of echo state networks known as Non-Linear Node with Delayed Feedback (NLDF) [17], [18]. In their work, they mention several types of faults produced by fuel cell stacks and only treat the cathode stoichiometric defects. Their work is based on supervised classification of labeled data, where the network's prediction is if a fault is occurring or not. Their scheme shows excellent classification rates in experimental studies, with 84% to 95% of accuracy. In a different application, Fan *et al.* [15] developed a fault detection scheme for the air compressors in city buses. They consider methods like echo state networks and a variation called Cycle Reservoir with Jumps (CRJ) [19]. In this case, the idea is to train two models with the same architecture, one with normal status and another with a failure present. Fault detection is achieved by a Consensus Self-Organized Models (COSMO) method to measure the different models' divergence. It was tested with two datasets: one synthetic and a real one. One key insight of their work is that Recurrent Neural Networks (RNNs), although effective in the synthetic data, does not provide a proper classification in the real one. A potential explanation for this is that their real data is noisy, making the echo state networks unable to learn the dynamics of the signal properly.

Also using an ANN architecture, Czajkowski and Patan [20] developed a fault detection strategy and applied it in a Twin Rotor Aero-Dynamical System (TRAS). Their approach is to use a combination of Leaky Echo State Networks and Model Error Modelling (MEM), providing a confidence region based on the residuals produced between real and estimated values. They show that this RNNs framework can successfully be used in diagnostic applications. Similarly, Westholm [16] uses echo state networks as a component of a process to detect a specific event on time-series obtained from electrical and mechanical systems. The focus is mainly on the *Feature Generation* component, using a Delay Line Reservoir architecture (DLR) [21]. Later, the approach is tested in three datasets named: Eyes, Occupancy, and Hard Drive. For each dataset, a specific architecture of the echo state network component is used. As a result, Eyes and Occupancy provide an F-measure of over 97%, whereas Hard Drive has under 16%. The explanation of this low rate comes from the fault detection policy and the short-term memory of

the RNN component defined in its architecture, which leads to early warnings.

B. ECHO STATE NETWORKS

As shown in the previous section, a specific ANN family, named *echo state networks* has been particularly useful in fault detection schemes. This usefulness is because, in many of the aforementioned applications, the studied system's behavior is usually nonlinear. Echo state networks (ESNs) are a type of recurrent neural network that has a dynamical memory to preserve in its internal state a nonlinear transformation of the input's history. Hence, they have shown to be exceedingly good at modeling nonlinear systems. Another advantage of ESNs is that they are easy to train because they do not need to backpropagate gradients as classical ANNs do.

An ESN can be defined as follows: consider a discrete-time neural networks like in [22]–[25], with N_u input units, N_x internal units (also called reservoir units), and N_y output units. Activations of input units at time step n are $\mathbf{u}(n) \in \mathcal{R}^{N_u}$, of internal units are $\mathbf{x}(n) \in \mathcal{R}^{N_x}$, and of output units $\mathbf{y}(n) \in \mathcal{R}^{N_y}$. The connection weight matrix $\mathbf{W}^{in} \in \mathcal{R}^{N_x \times (1+N_u)}$ for the input weights, $\mathbf{W} \in \mathcal{R}^{N_x \times N_x}$ for reservoir connections, $\mathbf{W}^{out} \in \mathcal{R}^{N_y \times (1+N_u+N_x)}$ for connections to the output units, and $\mathbf{W}^{fb} \in \mathcal{R}^{N_x \times N_y}$ for the connections that are projected back (also called feedback) from the output to the internal units. The connections go directly from input to output units and connections between output units are allowed. Fig. 1 shows the basic network architecture.

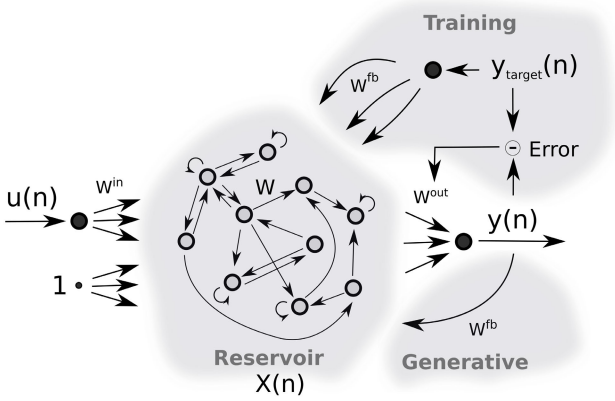


FIGURE 1. The basic echo state network architecture.

The activation of reservoir units are represented by

$$\tilde{\mathbf{x}}(t+1) = \tanh\left(\mathbf{W}^{\text{in}}\left[1; \mathbf{u}(t+1)\right] + \mathbf{W}\mathbf{x}(t) + \mathbf{W}^{\text{fb}}\mathbf{y}(t)\right), \tag{1}$$

and are updated according to

$$\mathbf{x}(t+1) = (1-\delta)\mathbf{x}(t) + \delta\tilde{\mathbf{x}}(t+1), \tag{2}$$

where $\delta \in (0, 1]$ is the leaky integrator rate. The output is calculated by

$$\mathbf{y}(t+1) = \mathbf{W}^{\text{out}}\left[1; \mathbf{u}(t+1); \mathbf{x}(t+1)\right], \tag{3}$$

where $[\cdot;\cdot]$ denotes the vertical vector concatenation. The coefficients in $\mathbf{W}^{\text{out}}$ are computed by using ridge regression, solving the following equation,

$$\mathbf{Y}_{target} = \mathbf{W}^{\text{out}}\mathbf{X}, \tag{4}$$

where $\mathbf{X} \in \mathcal{R}^{(1+N_u+N_x)\times T}$ with columns $[1; \mathbf{u}(t); \mathbf{x}(t)]$ for $n = 1, \ldots, T$; and all $\mathbf{x}(t)$ are produced by presenting the reservoir with $\mathbf{u}(t)$ and $\mathbf{Y}_{target} \in \mathcal{R}^{N_y \times T}$.

Finally, the solution can be represented by

$$\mathbf{W}^{out} = \mathbf{Y}_{target}\mathbf{X}^T\left(\mathbf{X}\mathbf{X}^T + \lambda\mathbf{I}\right), \tag{5}$$

where $\mathbf{I} \in \mathcal{R}^{(1+N_u+N_x)\times(1+N_u+N_x)}$ is the identity matrix and λ is a regularization factor (ridge constant). The ridge constant is estimated using grid search and time series cross-validation methods.

C. OUR CONTRIBUTION

Our work has the following novel contributions:

1) We develop a novel automatic fault detection scheme using an Echo State Networks as a component for dynamic learning.
2) We develop noise reduction techniques for highly noisy signals using evolutionary algorithms such as genetic algorithms and particle swarm optimization, improving the signal-to-noise ratio while maintaining its relevant characteristics such as trends and stationarity.
3) We develop a time-shift prediction process for comparison with the lower bound of a fault-free behavior.
4) We show that our approach has a good performance in noise real-life operational data from ALMA antennas, comparing the result to the real fault occurrences logged in the maintenance system.

As mentioned before, the main advantage of the proposed scheme over model-based techniques is that it does not require human intervention to develop or identify the relevant models. This fact reduces the time required for tuning the model, making it more scalable and applicable to other settings.

Our approach was tested on a specific type of slow degradation fault in the Intermediate Frequency Processors (IFPs) of the antennas of the ALMA observatory. This subsystem is critical to process the information gathered by each antenna and communicate it, reliably, to the correlator for processing. Due to the harsh environment in which the antennas operate, the operational data of the IFPs have a meager signal-to-noise ratio, making it very difficult for conventional fault detection methods to have an acceptable performance. Furthermore, the use of evolutionary algorithms and reservoir computing gives our approach a dynamic adaptation for a broader family of systems.

The rest of this article is organized as follows. First, in Section II, we detail the application problem, showing the main characteristics of the data used and the problem at hand. In Section III, we present the proposed fault detection scheme and explain how we train the system with real noisy data.

FIGURE 2. The antennas of the ALMA observatory.

Finally, in Section IV, we study the performance of our approach compared to the maintenance data from the IFPs, which is used as ground truth. We also include an Appendix, with the complete results of applying our scheme to more than 5 years of data from ALMA.

II. APPLICATION SETTING: THE ALMA TELESCOPE

The Atacama Large Millimeter/submillimeter Array (ALMA) is a revolutionary instrument operating in the very thin and dry air of northern Chile's Atacama desert, at an altitude of 5,200 meters above sea level. ALMA is one of the first industrial-scale new generation telescopes, composed of an array of 66 high-precision antennas working together at the millimeter and submillimeter wavelengths, corresponding to frequencies from about 30 to 950 GHz. Adding to the observatory's complexity, these 7 and 12-meter parabolic antennas, with extremely precise surfaces, can be moved around on the high altitude of the Chajnantor plateau to provide different array configurations, ranging in size from about 150 meters to up to 20 kilometers. The ALMA Observatory is an international partnership between Europe, North America, and Japan, in cooperation with the Republic of Chile [26].

ALMA is a very complex instrument. Each telescope is composed of hundreds of individual electronic and mechanical parts, each carefully calibrated, set up, and interconnected. This complexity is multiplied by 66, the number of single antennas, and the additional particularities contributed by the four distinct antenna designs developed. For some subsystems, the number of parts is duplicated, since two polarizations are being observed. Adding to this mix are the central equipment, the correlators, and the central local oscillator, which allow the whole array to perform as a single instrument through interferometry. Although not part of the telescope per se, ancillary or infrastructural systems, such as weather stations and power plants, are critical to attaining all the scientific objectives.

On top of the aforementioned technical intricacy, is the observatory's setting. The Chajnantor Plateau, with its perfect skies for astronomical observation, is also known for its extreme weather and oxygen-deprived air conditions that severely diminish troubleshooting and decision-making skills of human operators. Remote and automated tasks execution and investigation of arisen problems is a must, to the maximum possible extent [9]. Hence, developing automated systems that can reduce human intervention and detect possible failures ahead of time is extremely important.

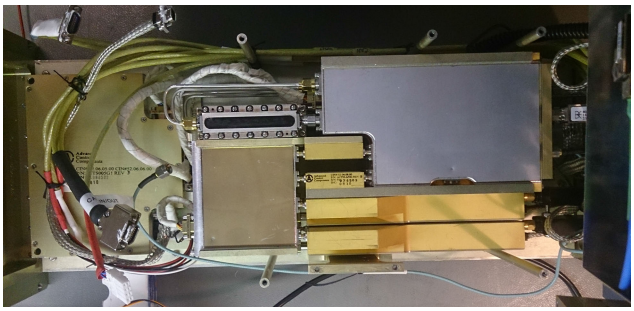


FIGURE 3. IFP module.

A. PROBLEM DESCRIPTION

The Intermediate Frequency Processor (IFP) of the antennas of the ALMA telescope, as described in [8], is a critical component responsible for the second down-conversion, signal filtering, and amplification of the total power measurement of sidebands and basebands. This subsystem allows for effective communication of the captured data to the central correlator for processing, thus making it a central and critical component of each antenna. Figure 3 shows the IFP module. There are 2 IFPs per antenna, one for each polarization, and each IFP has sensors measuring currents of three different voltage levels: 6.5, 8, and 10 volts. For 6.5 and 8 volts, currents have four different basebands: A, B, C, and D, whereas, for 10 volts, sidebands USB and LSB, and switch matrices SW1 and SW2 currents are read. Each current is sampled every 10 minutes. Figure 4 shows the IFP's currents for 6.5V,

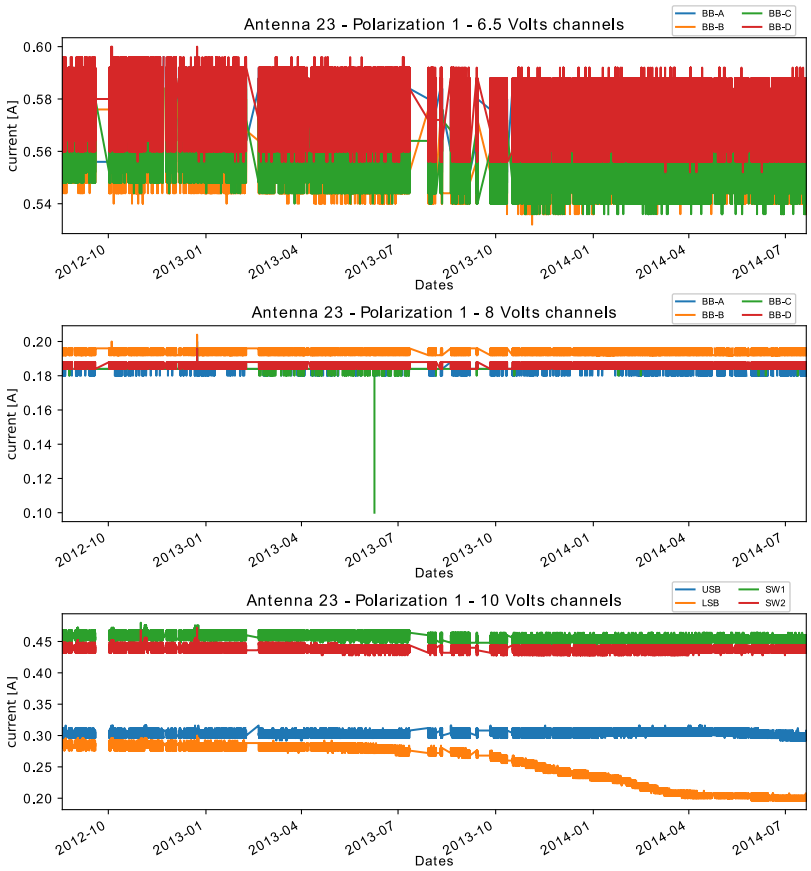


FIGURE 4. Currents from the IFP of one polarization of one antenna.

8V, and 10V. This example shows some typical characteristics of the raw data, with sections without samples, some currents present a high level of noise, there is redundant data, many different scales and large jumps in a single sample, records between channels have little time displacement and different numbers of samples, and sometimes outliers show up. Therefore, it is necessary to preprocess the data eliminating missing and duplicated records, as well as cleaning outliers.

One important thing to highlight from Figure 4 is that the signals are highly noisy, with high variance, and the ratio of mean to the standard deviation for each signal is very low, i.e., they have a low signal-to-noise ratio (SNR). Hence, identifying their trend is quite complicated. Because of this, as shown in Section III, we apply a preprocessing procedure. Given that the noise scale is different in every signal and IFP, we estimate the parameters of our procedure using an evolutionary algorithm [27]–[31], obtaining more effective parameter values for every IFP, and thus reducing as much as possible the noise present. The outcome of this preprocessing is a more suitable signal that maintains the relevant characteristics used later by our fault detection scheme.

B. DATA DESCRIPTION

The data exported from the ALMA operational database is stored in Hierarchical Data Format (HDF) files (.hdf or h5), used to store and organize large amounts of data.

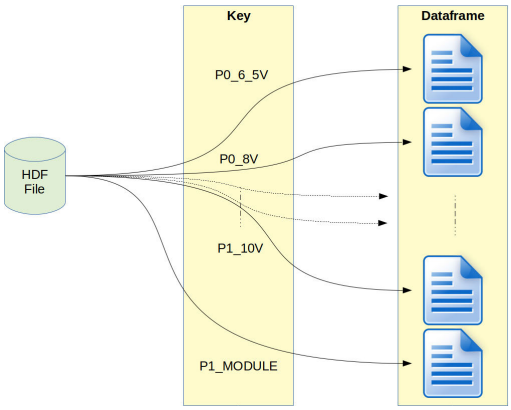


FIGURE 5. ALMA's HDF compressed file structure.

These files can be loaded in several programming languages like Python 3 (with Pandas Library) and R (with h5 Library). Each HDF file stores current data per antenna, polarization, and voltage levels, making a total of 528 HDF files. To make it more suitable, these files were compressed into 66 HDF files, one per antenna, such that each related dataset to an antenna was allocated in the file with a unique key, with a total of 8 keys per files (3 for volts per polarization and 2 for module serial number records). The structure and access to each dataset of the file are shown in Figure 5.

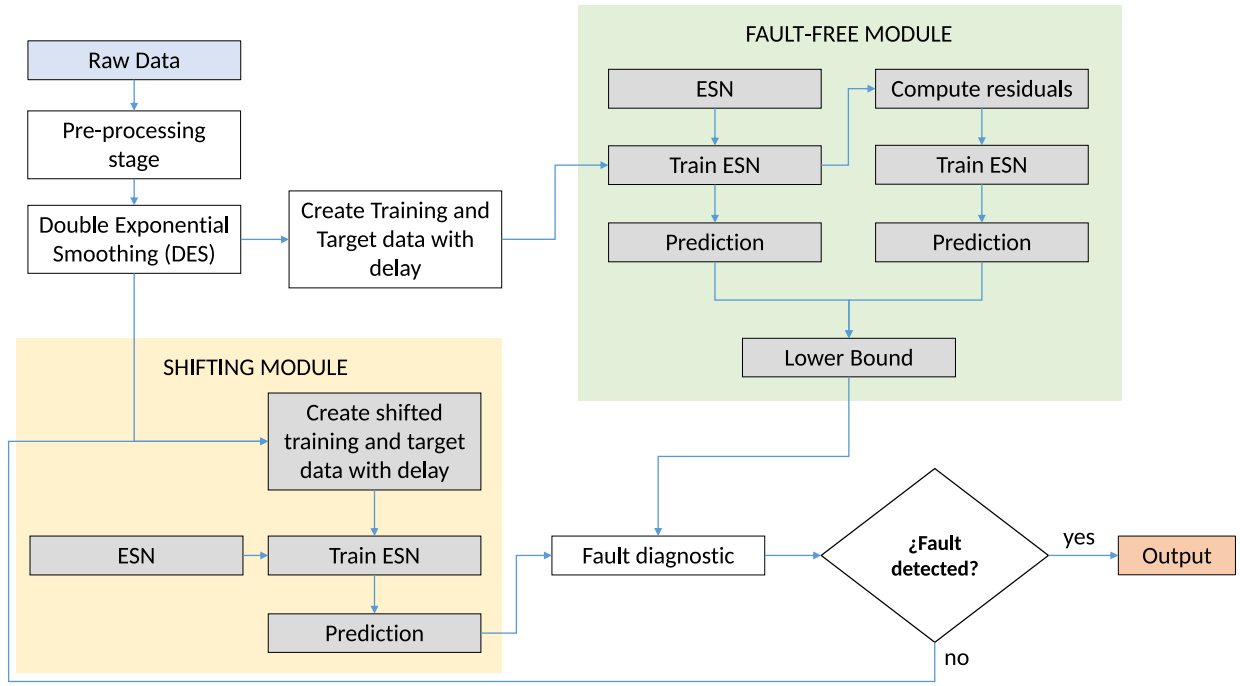


FIGURE 6. Fault detection scheme.

III. PROPOSED FAULT DETECTION SYSTEM

Our fault detection scheme was inspired by some of the ideas applied in [11], [20]. The complete fault detection process is summarized in Algorithm 3, and the whole process is showed in Figure 6. The following steps are the main ones of our fault detection scheme.

1) **Data pre-processing and cleaning**: the raw data is first reformatted to specific HDF dataframes as explained in II-B. Then, given that raw signals $\{y_t\}_{t=1}^{T}$ are highly noisy, have irregular time stamps, and possible outliers could be present, it is necessary to preprocess the signal to clean outliers, homogenize time between records, and finally, apply noise reduction. We use a double-exponential smoothing (DES) filter, tuned with evolutionary algorithms to generate a smoothed signal $\{u_t\}_{t=1}^{T}$.
l
2) **Fault-free characterization**: with the denoised signal, we compute a sequence of values that represent a lower bound of the fault-free prediction signal and will be used as a reference in the fault detection step. Such a lower bound is less sensitive to small variations in the signal.
3) **Time-shift forecast**: this step consists of making forecasts of the dynamics of the series using a time-shift strategy that allows us to use the information of a time block and obtain a more accurate prediction. This time-shift forecast is then used in the next step with lower bound jointly for the verification if a fault is present.
4) **Fault detection**: it is responsible for checking if there is a divergence between the time-shift forecast and the lower bound that will serve as an indicator that there is a failure. This fault detection process depends on a detection criterion that will allow us to reduce sensitivity in detecting premature cases and, therefore, possible false-positives cases.

The details of the previous steps are provided below.

A. DOUBLE EXPONENTIAL SMOOTHING PARAMETER ESTIMATION

The double exponential smoothing filter [9], [27]–[31] is a methodology used for forecasting in time series, but it can also be used for noise reduction. This method is particularly useful in time-series that have a trend, such as the current case-study, where the IFPs in faulty conditions present a slow degradation trend. This method depends on two parameters α and β, which need to be fixed appropriately. In order to learn the best values for these parameters, evolutionary algorithms are used. In particular, we apply a Genetic Algorithm and Particle Swarm Optimization as metaheuristics to determine the best values for α and β for each IFP. This fine-tuning is important since the different IFPs present different signal-to-noise ratios, and thus the preprocessing needs to be adjusted to each subsystem.

1) DOUBLE EXPONENTIAL SMOOTHING (DES)

Double Exponential Smoothing [9], [27], [28] is a technique used for forecasting in time series, also it can be used to

smooth time-series and perform noise reduction. It can be defined as follow:

Let $\{y_t\}$ be a sequence of observations beginning at time $t = 0$, suppose that $\{l_t\}$ represents the smoothed value, $\{b_t\}$ the best estimate of trend and $\{F_t\}$ the forecast at time t, then the formulae are given by:

$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{6}$$
$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \tag{7}$$
$$F_t = l_t + b_t \tag{8}$$

where $\alpha, \beta \in (0, 1)$, $b_0 = \frac{1}{N}\sum_{k=1}^{N}(y_{k+1} - y_k)$ and $l_0 = y_0$. In this method, α is the data smoothing factor, β is the trend smoothing factor; and for values close to zero higher the smoothing level.

2) GENETIC ALGORITHM (GA)

Genetic algorithms are a type of evolutionary algorithm introduced by John Holland in 1960, inspired by the process of natural selection [32]. This metaheuristic relies on operators such as mutation, crossover, and selection to find the best solutions of an estimation or optimization problem. The approach of using a Genetic Algorithm (GA) to estimate DES parameters has achieved a good approximation by minimizing Mean Absolute Error (MAE) as fitness function [27]–[29]. This function can be defined as:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|F_t - y_t|, \tag{9}$$

where $\{y_t\}$ is a sequence of observations, $\{F_t\}$ is a sequence of fitted forecast by DES, and T is the last observation time of the series.

We propose a variant of the approach for denoising each signal; it can be defined as follows.

Let $\{F_t^{\alpha}\}$ and $\{F_t^{1-\alpha}\}$ be the forecast sequences adjusted by DES using as parameters (α, β) and $(1-\alpha, \beta)$, respectively; given a constant $\tau \in [0, 1]$, we define the total weighted absolute error (TAE) as fitness function with the following expression:

$$TAE = \sum_{t=1}^{T}\tau\left|F_t^{\alpha} - y_t\right| + (1-\tau)\left|F_t^{1-\alpha} - y_t\right|. \tag{10}$$

The GA approach is summarized in Algorithm 1. The (α, β) values obtained by GA are used for forecasting; for smoothing $(1-\alpha, \beta)$ are used.

3) PARTICLE SWARM OPTIMIZATION (PSO)

Particle swarm optimization is another metaheuristic used for estimation and optimization. It is an evolutionary algorithm developed in 1995 by Kennedy and Eberhart [33], inspired by simulating social behavior and the observed movements of organisms such as insects, birds, and fish. This method has been applied in several optimization problems and gives high-quality results in a few iterations. PSO has also been used to estimate DES parameters in forecast applications [30], [31], by minimizing (9) as the fitness function. Given this insight, we apply our variant approach for denoising signals using (10) as a fitness function. The PSO process for denoising each signal is summarized in Algorithm 2.

Algorithm 1 DES - Genetic Algorithm

Require: Set a maximum evolution n, population size N, probability of mutation p, tournament size k and tolerance tol.

Ensure: Best denoise DES parameters value.

1: Randomly generates an initial population **P** of pairs $(\alpha, \beta) \in (0, 1)^2$, of size N.
2: **for** each individual in population **P do**
3: Compute individual's fitness values by using (10).
4: **end for**
5: $t = 0$,
6: **while** $(t < n)$ and ($\nexists$ fitness less than tol) **do**
7: Select k individuals with low fitness by tournament of size k for mating.
8: Produce a new set of individuals $\mathbf{P}'$ by uniform crossover and mutation with probability p.
9: **for** each individual in population $\mathbf{P}'$ **do**
10: Compute individual's fitness values by using (10).
11: **end for**
12: Add new population $\mathbf{P}'$ to **P**
13: Select top N individuals with lower fitness values.
14: $t = t + 1$.
15: **end while**
16: **return** $\{1-\alpha, \beta\}$, such that $\{\alpha, \beta\}$ has the lowest fitness.

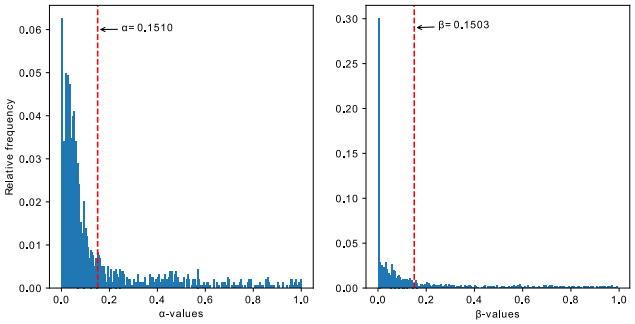


FIGURE 7. Histogram of DES parameters.

B. FAULT-FREE CHARACTERIZATION

Once the signal-to-noise ratio is improved in the previous step, the fault detection scheme characterizes the fault-free data. To do so, the process assumes that the subsystem starts in a fault-free mode, and uses that initial time-period, of length s, to learn the signal's characteristics in this setting, using an ESN. The ESN is trained using this initial data to predict later how the system behaves, and use it as a reference. We compare this reference with the real signal values to achieve fault detection, as we will explain later.

Algorithm 2 DES - PSO

Require: Set a maximum iteration n, number of particles N and tolerance tol.

Ensure: Best denoise DES parameters value.

Set the best global fitness value $s = 10^6$ and the best global position $\mathbf{g}$ randomly on the grid $(0, 1)^2$.
Initialize a set $\mathbf{P}$ of N particles.
for each particle in $\mathbf{P}$ **do**
    Initialize a random position $\{\alpha, \beta\}$ on the grid $(0, 1)^2$, random velocity in $[-1, 1]^2$ and set the best fitness value equal to 10^6 of the particle.
    Compute particle's fitness value by using (10)
    if fitness value is less than best fitness value **then**
        Update the best fitness value and best position of the particle
    end if
    if best fitness value is less than s **then**
        Update global fitness value s with the particle's best fitness value.
        Update global best position $\mathbf{g}$ with the particle's best position.
    end if
end for
$t = 0$,
while $(t < n)$ and $(s > tol)$ **do**
    for each particle in $\mathbf{P}$ **do**
        Update the particle's velocity given the global best position $\mathbf{g}$.
        Update the particle's position.
        Compute particle's fitness value by using (10).
        if fitness value is less than best fitness value **then**
            Update the best fitness value and best position of the particle.
        end if
        if best fitness value is less than s **then**
            Update global fitness value s with the particle's best fitness value.
            Update global best position $\mathbf{g}$ with the particle's best position.
        end if
    end for
    $t = t + 1$.
end while
return $\{1 - \alpha, \beta\}$, such that $\{\alpha, \beta\}$ has the lowest fitness.

Algorithm 3 Fault Detection Process

Require: signal: $\{y_t\}_{t=1:T}$, shift step: h, consecutive observations: N.

Ensure: Date of Fault detected

$flag \leftarrow False$
Apply denoising process to obtain $\{u_t\}_{t=1:T}$
Using $\{u_t\}_{t=1}^{s}(s < T)$, train an ESN with $\{u_t\}_{t=1}^{s'}$ $(s' < s)$, predict in generative mode to obtain $\{p_t\}_{t=s'+1}^{T}$
Compute residuals $\{r_t\}_{t=s'+1}^{s}$, train an ESN and predict $\{r_t\}_{t=s+1}^{T}$.
Compute lower bound $\{LB_t\}_{t=s+1}^{T}$ using (11).
$i \leftarrow 0$
while ¬flag **do**
    $i \leftarrow i + h$
    Train a ESN using $\{u_t\}_{t=i}^{s+h}$ and predict $\{p_t\}_{t=i+1}^{T}$.
    if $\exists t$ such that $p_t < LB_t$ for N consecutive observations **then**
        $flag \leftarrow True$
    end if
end while
return date belonging to position i.

Although the signal should remain stable, the real data shows that it still has some variability level, which could generate false positives in the fault detection step. To reduce this non-desired effect, we define a variation threshold by analyzing the residuals of the prediction, and use this threshold as a lower bound as follows.

Consider the first s values $\{u_t\}_{t=1}^{s}$ $(s < T)$, lets train an ESN with $\{u_t\}_{t=1}^{s'}$ $(s' < s)$, and predict in a generative mode to obtain a signal $\{p_t\}_{t=s'+1}^{T}$. We then compute the residual as $\{r_t = u_t - p_t\}_{t=s'+1}^{s}$ that will be fed to another ESN to fit and predict $\{r_t\}_{t=s+1}^{T}$. Finally, we define the lower bound as a combination of the free-fault and residual predictions expressed as the following,

$$LB = \{LB_t\}_{t=s+1}^{T} = \{p_t - k \cdot r_t\}_{t=s+1}^{T}, \tag{11}$$

where LB is another series that behaves as a forecast of fault-free tolerance through time, and $k > 0$ a gap constant. The sequence of steps to these calculations are summarized in a diagram shown in the Figure 8.

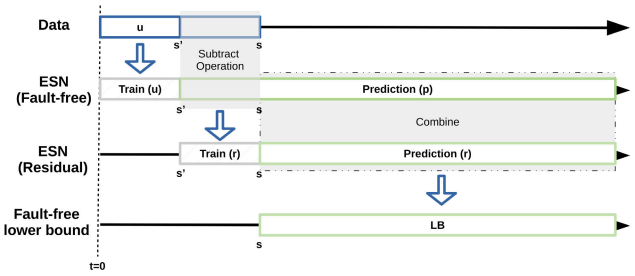


FIGURE 8. Fault-free lower bound diagram.

C. TIME-SHIFT FORECAST STAGE

It is relevant to highlight that the exact moment where the fault occurs and the subsystem begins to degrade is unknown. Unlike most of the fault detection literature, where the system transitions from an unfaulty state to a faulty one relatively fast, this transition can be extremely lengthy in the case of degradation faults. This slow transition is specifically challenging for conventional pattern recognition techniques. Furthermore, since there is no labeled data identifying when

the system is faulty, a supervised learning technique cannot be used to learn the system's dynamics under a faulty condition.

To overcome the previously mentioned limitations, we developed a learning and prediction scheme based on an ESN, which is time-shifted to compare the changes in the system's dynamics. This technique also reduces the need for doing feature engineering manually by analyzing the data beforehand.

Figure 9 gives a general idea on our approach. The prediction of the ESN is shifted in time by h units, and its prediction is compared to the fault-free lower bound to identify the current operation mode. Hence, given a fixed $h > 0$, we take a sequence $\{u_t\}_{t=h}^{s+h}$ to train a new ESN and compute an element-wise average prediction for n repetitions, obtaining $\{p_t\}_{t=s+h+1}^{T}$. This signal will be used to verify if fault is present in the fault detection stage.

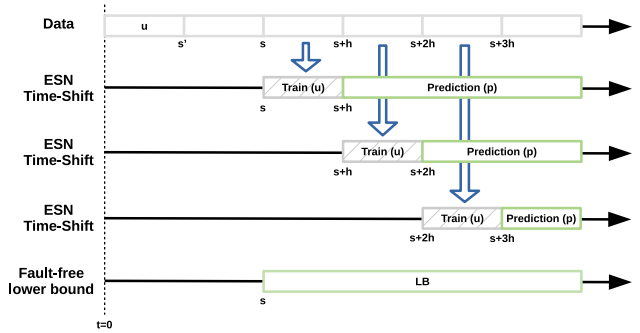


FIGURE 9. Time-shift forecast diagram.

D. FAULT DETECTION STAGE

The fault detection stage is responsible for verifying whether there is a point in which the time-shift prediction diverges from the computed lower bound. Defining the procedure to signal a detected fault when there is a divergence between the two signals is key to managing our scheme's performance.

A criterion that is not sensitive enough can cause late detections or false negatives. This behavior is especially troublesome in our application: the occurrence of a fault leads to disabling one of the antennas of the array, which in some cases implies the whole array cannot function as needed. Furthermore, to fix it, technicians need to work in a harsh environment given the location of the antennas.

On the other hand, indicating a fault at the first sign of divergence can lead to many false-positives. The cost of this error is lower than the false negatives since it will only trigger an engineer to check the data and confirm the fault. Still, a high number of false-positives will lead to the maintenance losing valuable time reviewing data.

In our scheme, we balance these two errors by using a simple criterion based on the count of consecutive cases in which the prediction is lower than the lower bound. If, in a time segment, the comparison reaches N consecutive cases, then it will be indicated that it found a fault and report the time when it happened, i.e., a fault is detected at time t.

$$p_k < LB_k, \quad k = t - N, \ldots, t. \tag{12}$$

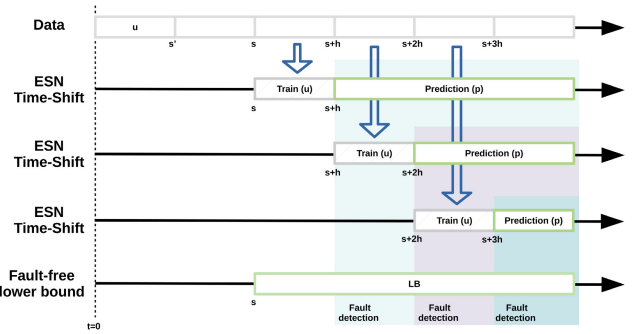


FIGURE 10. Fault detection stage diagram.

Otherwise, it will return to the previous step (time-shift forecast) and compute the next time-shift ESN (increasing the displacement by h) and compare its forecast with the lower bound. This procedure is repeated until reaching the end of the data stream. Figure 10 shows a diagram on how the fault detection and the time-shift procedures interact to determine when a fault is detected.

IV. IMPLEMENTATION AND RESULTS

To evaluate its performance, we tested our fault detection scheme with real historical data. As explained before, we used seven years of operational data, from 2012 to 2019, from all 132 IFPs at the ALMA radio telescope. The data was first cleaned and organized, as detailed in Section II-B.

For the implementation, the ESN parameters were as follows. We used three input units (two delays included), one output unit, 500 reservoir units, 40% sparsity rates, a spectral radius of 0.995, and 0.1 as the leaky rate. We used 4,500 observations for training and 4,500 observation for the time-shift.

For the GA, we used a maximum of 20 iterations, with a 1.2 tolerance value. The mating group size was fixed to 10, with ten tournaments. The population size was set to 100, with a 50% probability for mutations.

For the PSO, we set to 5 the maximum number of iterations, with eight particles, and 0.5 as the inertia constant. Both the social and cognitive constants were set to 1. We also set $\tau = 0.6$ in the *TAE* function.

Finally, for our fault detection process, we set k to 4, with a shifting of 4,500 (which represents about one month of observations). At least 10,000 observations belonging to a single IFP serial number were needed (about two months of data) in case different serial numbers had different noise profiles. Considering that the fault can be present for more than a year before problems happen in the system, this initial requirement does not limit the application significantly.

All algorithms were implemented in Python 3.7 and ran on a computer with an Intel(R) Core(TM) Processor i7-3770S of 3.10 GHz, with 8 GB RAM, and using Windows 7 SP1 (64 bits) as OS.

Table 1 shows the parameters estimated by the GA and PSO methods used in the DES filter for denoising signals.

TABLE 1. Example of DES parameters for smoothing each baseband (Polarization 1, 6.5V channels) and runtime [sec] for antenna 23.

Method	Channel	α	β	I_{rate}	runtime
GA	BB-A	0.0812	0.1145	2.39	84.30
	BB-B	0.0825	0.1296	2.39	84.32
	BB-C	0.0704	0.0678	2.32	84.52
	BB-D	0.0207	0.0367	3.60	84.49
PSO	BB-A	0.0485	0	3.29	17.49
	BB-B	0.1795	0	2.51	17.69
	BB-C	0.0377	0.0356	2.56	17.48
	BB-D	0.0190	0.0248	3.81	17.41

The table shows the particular case of the data for Antenna 23 and Polarization 1, but the results were similar across all antennas. Figure 13 shows the noise reduction achieved using PSO.

From Table 1, we can see that a different set of smoothing parameters (α, β) for the DES filter are determined for every different baseband. They also differ depending on the methodology used to compute them, GA or PSO. Although the performance in denoising the signals was similar for both methodologies, as Table 1 shows, the time required by PSO was vastly superior. Therefore, combining PSO with DES was deemed useful for signals with different noise scale.

It is also important to highlight that the estimated parameters are close to zero in most cases, as illustrated in Figure 7, which shows a histogram of all the 1,584 estimations made. As expected, these values imply that signals with a high level of noise need a higher smoothing level (i.e., parameter values close to zero). This parameter, in turn, reduces the variability in the signal while smoothing the slope between observations. Our approach is successful in increasing the signal-to-noise ratio, as shown in Table 1. We computed how much this ratio increases between the original signal and the filtered one, shown as I_{rate} in Table 1.

As shown in Figure 7, in most cases the value of parameter β is close to zero. Since it is fairly similar in most cases, we studied the performance of our filtering stage when neither PSO nor GA is used for tuning. In this case, we set $\beta = 0$, and test for a few values of $\alpha \in [0.05, 0.1]$. Our results show that the filtering keeps on delivering good quality output of the denoised signal. The best performance was achieved for $\alpha = 0.1$, which was the value later used in our fault detection scheme when testing over all our data.

TABLE 2. Fault detection and runtime [sec] for antenna 23.

Polarization	Volts	Channel	Fault date	runtime
0	6.5	BB-B	2014-03-10	208.99
0	6.5	BB-D	2014-08-02	176.37
0	10	SW1	2013-12-08	198.24
1	10	LSB	2013-12-05	145.88

The results of running our fault detection scheme for one antenna are shown in Table 2 and illustrated in Figure 14.

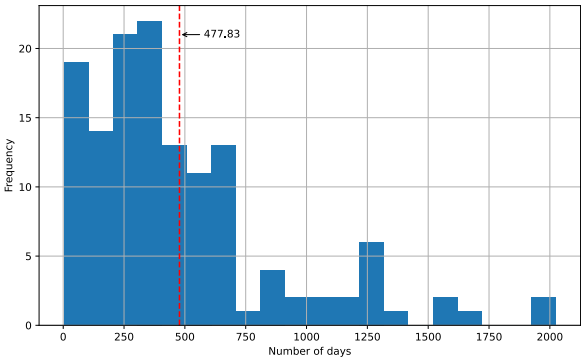


FIGURE 11. Histogram of the number of days that our FDD Scheme detects a fault before the human operator.

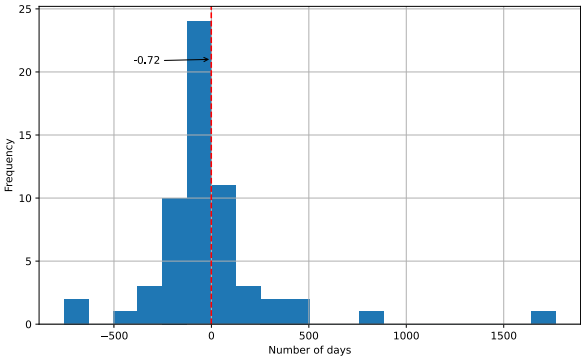


FIGURE 12. Histogram of the number of days the ESN detects a fault before the FDD Scheme in [8].

TABLE 3. Detection improvement.

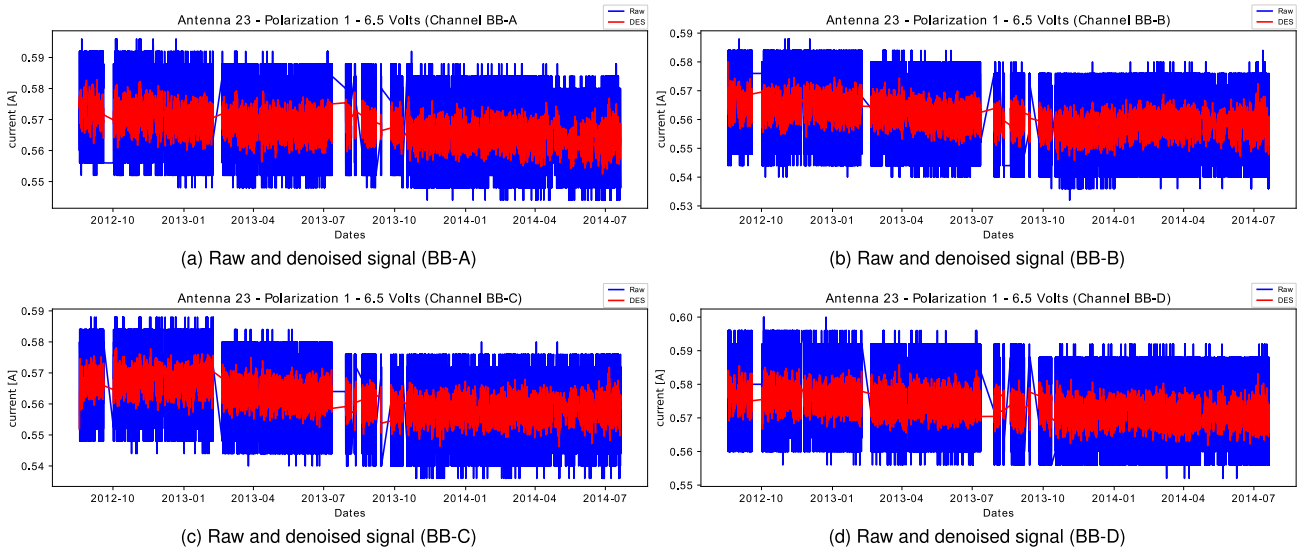
Number of months	≤ 6	≤ 12	≤ 24	≤ 36	Total
Number of cases	33	69	113	121	136

TABLE 4. Confusion matrix of the fault detection scheme.

		Ground truth	
		Fault	No fault
ESN	Fault	90 (52%)	46 (27%)
	No fault	6 (3%)	31 (18%)

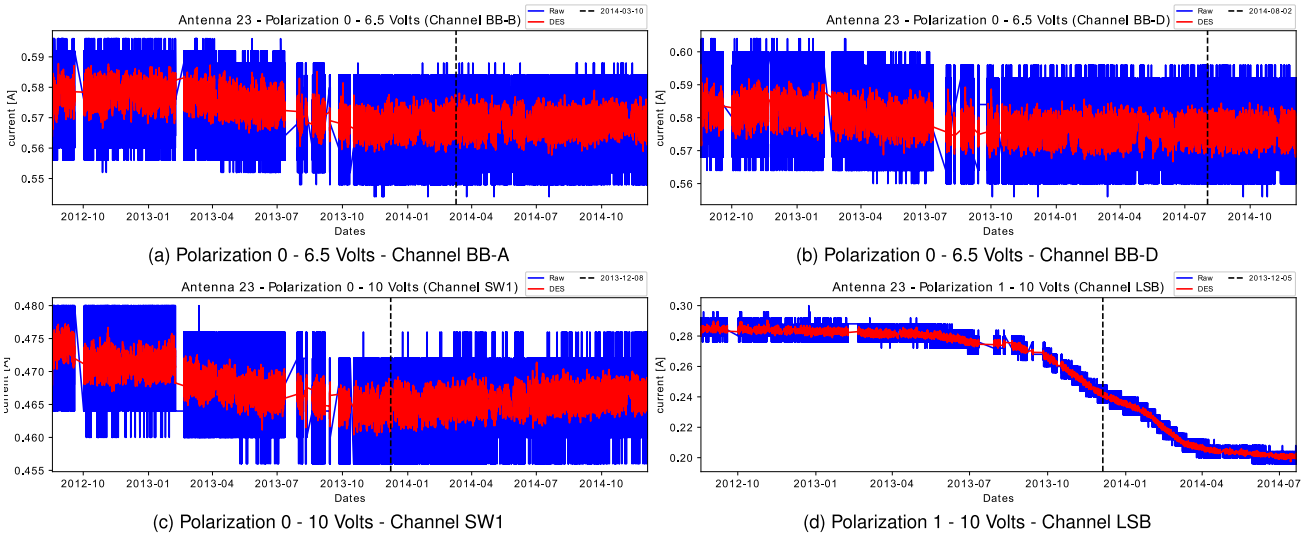
Table 2 shows the results for the two IFPs in the antenna (one for each polarization), showing that a fault can occur on each of them at different times. This example is also useful since it depicts that a fault can manifest by affecting only a single channel (like in the Polarization 1 IFP, where only the LSB channel was affected), or many (like the Polarization 0 IFP, where three different channels were affected). Moreover, when many channels are affected, the dates at which each channel can present the effects can vary significantly.

Figure 14 shows that each fo the different channels of each voltage can have vastly different dynamics during the faulty-operation phase. In the Polarization 0 IFP, the 10-volt SW1 channel has a higher degradation level and is thus detected first. The rest, having a slower degradation, are detected much later. In Figure 14, we can also appreciate



(a) Raw and denoised signal (BB-A)
(b) Raw and denoised signal (BB-B)
(c) Raw and denoised signal (BB-C)
(d) Raw and denoised signal (BB-D)

FIGURE 13. Application of DES using PSO. parameters in Table 1.



(a) Polarization 0 - 6.5 Volts - Channel BB-A
(b) Polarization 0 - 6.5 Volts - Channel BB-D
(c) Polarization 0 - 10 Volts - Channel SW1
(d) Polarization 1 - 10 Volts - Channel LSB

FIGURE 14. Raw and DES signal with fault date - Antenna 23.

the effect of the noise reduction stage, which significantly improves the information in the signal. In these cases, where a fault is detected through multiple channels, we consider the earliest detection to be the fault detection date for that IFP. In the example of Table 2, the IFP of Polarization 0 in Antenna 23 would have assigned as fault date 2013-12-08.

To measure our scheme's effectiveness, we compared the results of our system with a *Ground Truth*. This ground truth was constructed from the original analysis in [8], later validated by the engineering team at ALMA by manually revising the maintenance logs, to determine the actual date when the IFP was detected to be faulty by a human operator. A summary of all the analyzed cases is shown in Appendix. In Tables 5 and 6 we give a general overview of all the faults detected in the data from 2012 to 2019, and compare them to the ground truth identified by the engineering personnel.

In total, there were 173 faults detected by our scheme. In some cases, the same IFP could fail several times within the large time-window analyzed. Hence, an antenna and polarization could appear more than once in the table. In the *Gap* column of these tables, we show the time difference (in days) between the detection time of our system and when the engineering team validated that the subsystem was in faulty mode. That detection took place, in most cases, when the IFP stopped working entirely and was replaced. Table 3 summarizes this gap information, showing the number of detections made prior to what was set in our ground truth data. Figure 11 details our results further, showing a histogram with the number of days gained by our scheme, compared to when the human operators detected the fault. In average, our scheme could detect a fault 477.8 days earlier, and in some cases detection was even a couple of years earlier. These cases

TABLE 5. Fault detection performance for antennas 01 to 41.

Antenna	Polarization	Ground Truth[1]	ESN	Δ^2	Antenna	Polarization	Ground Truth[1]	ESN	Δ^2
01	0	2013-Aug-04	2012-Oct-08	300	22	0	2017-May-03	2013-Oct-17	1294
01	0	2015-Feb-19	2014-Dec-04	77	22	1	2019-Mar-04	2014-Dec-06	1549
01	0	2017-Mar-31	2016-Dec-20	101	23	0	2014-Dec-04	2013-Dec-20	349
01	1	No fault	No fault	.	23	0	2019-Mar-18	2018-May-23	299
02	0	2013-Jul-10	2013-Mar-14	118	23	1	2014-Jul-20	2013-Dec-05	227
02	1	2014-Sep-23	2013-Mar-20	552	24	0	2014-May-11	2014-Feb-13	87
03	0	2019-Mar-18	2013-Sep-17	2008	24	0	2019-Mar-18	2017-Oct-07	527
03	1	2019-Mar-18	2015-Oct-10	1255	24	1	2014-May-11	2014-Mar-17	55
04	0	2014-Nov-02	2013-Apr-19	562	25	0	2013-Oct-27	2013-Sep-29	28
04	1	2014-Nov-02	2014-Feb-02	273	25	1	2015-Sep-10	2015-May-02	131
05	0	2014-Mar-21	2013-Mar-30	356	26	0	2016-Oct-10	2015-Mar-25	565
05	0	2018-Apr-13	2018-Mar-15	29	26	1	2015-Jul-03	2014-Feb-12	506
05	1	2014-Mar-21	2013-Mar-30	356	27	0	2014-Aug-31	2013-Oct-10	325
06	0	2014-Oct-25	2013-Aug-13	438	27	1	2014-Jul-20	2013-Mar-30	477
06	0	2018-Apr-13	2018-Mar-10	34	27	1	2015-Jun-08	2014-Nov-14	206
06	1	2019-Mar-18	2018-May-05	317	27	1	2019-Mar-18	2017-Oct-27	507
07	0	2015-Jan-17	2013-Nov-19	424	28	0	2015-Jan-01	No fault	.
07	1	2015-Jan-17	2014-Apr-02	290	28	1	2015-Jan-01	2014-Apr-02	274
08	0	2014-Jun-05	2013-Dec-29	158	28	1	2019-Mar-04	No fault	.
08	1	2014-Jun-06	2013-Dec-28	160	29	0	2015-Apr-27	2015-Apr-22	5
09	0	2017-May-21	2017-May-07	14	29	0	2018-Apr-20	2016-May-18	702
09	1	2014-Sep-30	2013-Dec-13	291	29	1	2014-Jan-19	No fault	.
09	1	2017-Jan-03	2016-Jul-13	174	30	0	2019-Mar-18	2018-Jul-31	230
10	0	2013-Dec-27	2013-Sep-20	98	30	1	2014-Mar-19	No fault	.
10	1	2013-Dec-27	2013-Mar-23	279	30	1	2015-Jun-04	2015-Feb-10	114
11	0	2017-Mar-24	2014-May-20	1039	31	0	No fault	No fault	.
11	1	2019-Mar-08	2018-Feb-21	380	31	1	2017-Nov-16	2017-Jan-22	298
11	1	2017-Mar-24	2014-Sep-16	920	32	0	2016-Oct-21	2016-Mar-11	224
12	0	2019-Jan-11	2018-Nov-27	45	32	1	2015-Aug-29	2014-Aug-08	386
12	1	No fault	No fault	.	33	0	No fault	No fault	.
13	0	2014-Sep-09	2013-Oct-07	337	33	1	No fault	No fault	.
13	1	2014-Sep-09	2012-Oct-25	684	34	0	No fault	No fault	.
13	1	2017-Sep-11	2017-Mar-30	165	34	1	No fault	No fault	.
14	0	2019-Mar-18	2016-Oct-16	883	35	0	2015-May-06	2014-May-20	351
14	1	2014-Jun-04	2013-Aug-22	286	35	0	2019-Mar-04	2016-May-31	1007
14	1	2019-Mar-18	2017-Dec-18	455	35	1	2015-Dec-11	2014-Dec-28	348
15	0	No fault	No fault	.	36	0	2019-Mar-18	No fault	.
15	1	2013-Apr-25	2012-Nov-03	173	36	1	2018-Aug-07	2018-Jun-11	57
16	0	No fault	No fault	.	37	0	No fault	No fault	.
16	1	2015-Apr-01	2013-Jun-30	640	37	1	No fault	No fault	.
16	1	2019-Mar-18	2015-Sep-06	1289	38	0	2015-Oct-22	2014-Nov-10	346
17	0	2014-Nov-02	No fault	.	38	1	No fault	No fault	.
17	1	2016-Nov-08	2013-Oct-27	1108	39	0	2014-Nov-26	2013-Feb-05	659
18	0	No fault	No fault	.	39	0	2017-Mar-15	2015-Dec-17	454
18	1	2015-Apr-14	2013-Nov-08	522	39	1	2016-Apr-22	2014-Aug-10	621
18	1	2016-Jul-14	2016-Jun-15	29	40	0	No fault	No fault	.
19	1	2016-Jul-14	2016-May-21	54	40	1	2014-Dec-06	2014-Jan-23	317
20	0	2015-Apr-22	2012-Nov-09	894	40	1	2017-Mar-09	2016-Sep-17	173
20	0	2019-Mar-18	2017-May-25	662	41	0	2014-Jun-24	2013-Feb-21	488
20	1	2015-Apr-22	2014-Feb-07	439	41	0	2015-Aug-14	2014-Oct-10	308
21	0	2019-Mar-04	2017-Oct-03	517	41	1	2015-Jun-19	2013-Jun-16	733
21	1	2013-Jun-18	2012-Nov-20	210	41	1	2018-Nov-26	2017-May-13	562
21	1	2019-Mar-04	2015-Oct-22	1229					

[1] The date of change of the IFP module.
[2] Number of days the detection is made before the Ground Truth date.

TABLE 6. Fault detection performance for antennas 42 to 66.

Antenna	Polarization	Ground Truth[1]	ESN	Δ^2	Antenna	Polarization	Ground Truth[1]	ESN	Δ^2
42	0	2014-Dec-13	2013-Mar-02	651	55	0	2014-Mar-04	2013-Feb-20	377
42	1	2016-Oct-10	2015-Feb-15	603	55	1	2014-Oct-25	2013-Aug-18	433
43	1	2015-Feb-19	2014-Jul-25	209	55	1	2017-Aug-18	2015-Nov-18	639
44	0	No fault	No fault	.	55	1	2019-Mar-18	2018-Apr-25	327
44	1	No fault	No fault	.	56	0	2019-Mar-18	2019-Mar-15	3
45	0	2019-Mar-18	2014-Jul-13	1709	56	1	2019-Mar-18	2013-Aug-31	2025
45	1	2014-Mar-02	2013-Mar-08	359	57	0	2014-Aug-14	2012-Nov-13	639
45	1	2019-Mar-18	2018-Jun-12	279	57	0	2019-Mar-18	2018-May-24	298
46	0	2015-Apr-01	2014-Aug-11	233	57	1	No fault	No fault	.
46	0	2017-Dec-01	2016-Apr-23	587	58	0	2016-Dec-03	2013-Nov-02	1127
46	1	2013-Dec-24	2013-Nov-30	24	58	1	2014-Dec-27	2013-Apr-13	623
46	1	2017-Dec-01	No fault	.	58	1	2016-Aug-13	2015-Dec-04	253
47	0	2019-Mar-04	No fault	.	59	0	2018-Nov-05	2018-Nov-03	2
47	1	2017-Oct-22	2016-Feb-16	614	59	1	No fault	No fault	.
48	0	2017-Jan-18	2014-Oct-29	812	60	0	No fault	No fault	.
48	1	2015-Jun-19	2014-Jun-08	376	60	1	2014-Nov-19	2014-May-23	180
48	1	2019-Mar-18	No fault	.	60	1	2019-Mar-18	2015-Oct-08	1257
49	0	2017-Feb-10	2016-Jan-05	402	61	0	2014-Sep-09	2013-Apr-18	509
49	0	2019-Mar-18	2018-Jan-20	422	61	0	2015-Sep-25	2015-Jul-16	71
49	1	No fault	No fault	.	61	1	2014-Sep-09	2014-Apr-18	144
50	0	2019-Mar-18	2015-Jun-25	1362	62	0	2017-Jun-24	2014-Mar-15	1197
50	1	No fault	No fault	.	62	0	2019-Mar-18	2018-Feb-27	384
51	0	2014-Dec-13	No fault	.	62	1	2014-Jun-06	2013-Sep-04	275
51	0	2019-Mar-18	2018-May-05	317	62	1	2016-Sep-29	2015-Aug-28	398
51	1	2015-Mar-08	2013-Aug-20	565	63	0	2017-Nov-20	2017-Jul-17	126
52	0	2015-May-28	2013-Jun-20	707	63	1	No fault	No fault	.
52	0	2017-Jul-14	2016-Nov-26	230	64	0	2014-Sep-14	No fault	.
52	1	2013-Apr-13	2012-Dec-23	111	64	1	2013-Apr-27	2012-Dec-26	122
52	1	2013-Nov-04	2013-Sep-04	61	64	1	2014-Sep-14	No fault	.
52	1	2017-Mar-08	2016-Jan-05	428	65	0	No fault	No fault	.
53	0	No fault	No fault	.	65	1	2019-Mar-18	2018-Mar-25	358
53	1	2015-Sep-11	2014-May-16	483	66	0	2015-Sep-25	2014-Jan-10	623
54	0	2017-Jan-04	2013-Jun-28	1286	66	1	2017-Jun-03	2013-Mar-29	1527
54	1	2015-Sep-25	2013-Jun-22	825					

[1] The date of change of the IFP module.
[2] Number of days the detection is made before the Ground Truth date.

were confirmed by the ALMA engineering team. The reason behind these few cases, was because the maintenance team, without realizing the IFP was starting to fail, improved communication gain in another subsystem, effectively retarding the total failure of the antenna.

Table 4 shows the confusion matrix summarizing the performance of our scheme. The two main performance metrics we can compute from this matrix are the *accuracy* (indicating how many of the cases were correctly identified), which is at 70%, and the *F-1 Score* (which is the harmonic mean of precision and sensitivity), which is at 78%. Both of these metrics are an important improvement compared to the currently implemented scheme. Additionally, from Table 4 we obtain has a recall rate of 94% and a precision of 66%. These results are promising as they show a meager number of false negatives (only 3%), which is crucial in our application. False-negatives imply that the array stops working as needed, and the maintenance teams need to go to the site to make repairs. Although the percentage of false-positives is not as low (at 27%), the effect is not as negative since it only implies some additional time by the engineering team to review the data and realize it is a false-positive. There is no need to go to the site to do maintenance procedures or measurements to validate the fault.

We also compared our scheme with the one currently in use by ALMA, and detailed in [8]. Figure 12 shows a histogram with the days gained by the ESN-based scheme. Although the reaction time is almost the same (in average 0.7 days slower), the scheme presented in [8] required that the maintenance team to determine the system's dynamics both in faulty and fault-free conditions. In our case, the ESN takes care of identifying the system's dynamics, making the process simpler, and easy to scale and apply in other settings. Additionally, Tables 7 and 8 give exhaustive details on how the current

TABLE 7. Fault detection comparison for antennas 01 to 41.

Antenna	Polarization	Ground Truth[1]	Kalman[3]	Δ^2	Antenna	Polarization	Ground Truth[1]	Kalman[3]	Δ^2
01	0	2013-Aug-04	2012-Sep-08	330	22	0	2017-May-03	No fault	.
01	0	2015-Feb-19	2014-Oct-29	113	22	1	2019-Mar-04	2012-Nov-08	2307
01	0	2017-Mar-31	No fault	.	23	0	2014-Dec-04	No fault	.
01	1	No fault	No fault	.	23	0	2019-Mar-18	2017-Aug-24	571
02	0	2013-Jul-10	No fault	.	23	1	2014-Jul-20	2013-Jul-25	360
02	1	2014-Sep-23	No fault	.	24	0	2014-May-11	No fault	.
03	0	2019-Mar-18	2018-Jul-21	240	24	0	2019-Mar-18	No fault	.
03	1	2019-Mar-18	No fault	.	24	1	2014-May-11	2014-Apr-08	33
04	0	2014-Nov-02	No fault	.	25	0	2013-Oct-27	No fault	.
04	1	2014-Nov-02	2013-Aug-11	448	25	1	2015-Sep-10	No fault	.
05	0	2014-Mar-21	2012-Dec-14	462	26	0	2016-Oct-10	No fault	.
05	0	2018-Apr-13	No fault	.	26	1	2015-Jul-03	2014-May-29	400
05	1	2014-Mar-21	2012-Aug-18	580	27	0	2014-Aug-31	No fault	.
06	0	2014-Oct-25	2013-Oct-10	380	27	1	2014-Jul-20	2013-Mar-22	485
06	0	2018-Apr-13	No fault	.	27	1	2015-Jun-08	2014-Oct-16	235
06	1	2019-Mar-18	No fault	.	27	1	2019-Mar-18	No fault	.
07	0	2015-Jan-17	No fault	.	28	0	2015-Jan-01	2013-Nov-15	412
07	1	2015-Jan-17	No fault	.	28	1	2015-Jan-01	2013-Nov-02	425
08	0	2014-Jun-05	No fault	.	28	1	2019-Mar-04	2017-Aug-20	561
08	1	2014-Jun-06	No fault	.	29	0	2015-Apr-27	No fault	.
09	0	2017-May-21	2017-Mar-21	61	29	0	2018-Apr-20	2017-Aug-22	241
09	1	2014-Sep-30	No fault	.	29	1	2014-Jan-19	2013-Jun-28	205
09	1	2017-Jan-03	2016-Sep-14	111	30	0	2019-Mar-18	2018-Jun-23	268
10	0	2013-Dec-27	No fault	.	30	1	2014-Mar-19	2013-Nov-16	123
10	1	2013-Dec-27	No fault	.	30	1	2015-Jun-04	No fault	.
11	0	2017-Mar-24	No fault	.	31	0	No fault	No fault	.
11	1	2019-Mar-08	No fault	.	31	1	2017-Nov-16	No fault	.
11	1	2017-Mar-24	No fault	.	32	0	2016-Oct-21	2015-Jul-08	471
12	0	2019-Jan-11	No fault	.	32	1	2015-Aug-29	2014-Aug-01	393
12	1	No fault	No fault	.	33	0	No fault	No fault	.
13	0	2014-Sep-09	2013-Oct-23	321	33	1	No fault	No fault	.
13	1	2014-Sep-09	2012-May-02	860	34	0	No fault	No fault	.
13	1	2017-Sep-11	No fault	.	34	1	No fault	No fault	.
14	0	2019-Mar-18	No fault	.	35	0	2015-May-06	2014-May-31	340
14	1	2014-Jun-04	No fault	.	35	0	2019-Mar-04	No fault	.
14	1	2019-Mar-18	2018-Oct-26	143	35	1	2015-Dec-11	2014-Oct-18	419
15	0	No fault	No fault	.	36	0	2019-Mar-18	2017-May-21	666
15	1	2013-Apr-25	No fault	.	36	1	2018-Aug-07	2018-Apr-16	113
16	0	No fault	No fault	.	37	0	No fault	No fault	.
16	1	2015-Apr-01	No fault	.	37	1	No fault	No fault	.
16	1	2019-Mar-18	No fault	.	38	0	2015-Oct-22	No fault	.
17	0	2014-Nov-02	2014-Jun-03	152	38	1	No fault	No fault	.
17	1	2016-Nov-08	No fault	.	39	0	2014-Nov-26	2012-Nov-01	755
18	0	No fault	No fault	.	39	0	2017-Mar-15	No fault	.
18	1	2015-Apr-14	2013-Oct-01	560	39	1	2016-Apr-22	2012-Nov-01	1268
18	1	2016-Jul-14	No fault	.	40	0	No fault	No fault	.
19	1	2016-Jul-14	2016-Apr-04	101	40	1	2014-Dec-06	No fault	.
20	0	2015-Apr-22	2012-Nov-02	901	40	1	2017-Mar-09	2015-May-11	668
20	0	2019-Mar-18	No fault	.	41	0	2014-Jun-24	2012-Jun-24	730
20	1	2015-Apr-22	2014-Jan-25	452	41	0	2015-Aug-14	2014-Oct-13	305
21	0	2019-Mar-04	No fault	.	41	1	2015-Jun-19	2014-May-05	410
21	1	2013-Jun-18	No fault	.	41	1	2018-Nov-26	2016-Nov-21	735
21	1	2019-Mar-04	No fault	.					

[1] The date of change of the IFP module.
[2] Number of days the detection is made before the Ground Truth date.
[3] FDD Scheme in [8]

TABLE 8. Fault detection comparison for antennas 42 to 66.

Antenna	Polarization	Ground Truth[1]	Kalman[3]	Δ^2	Antenna	Polarization	Ground Truth[1]	Kalman[3]	Δ^2
42	0	2014-Dec-13	2012-Oct-24	780	55	0	2014-Mar-04	2013-Jul-25	222
42	1	2016-Oct-10	No fault	.	55	1	2014-Oct-25	No fault	.
43	1	2015-Feb-19	2014-Nov-15	96	55	1	2017-Aug-18	No fault	.
44	0	No fault	No fault	.	55	1	2019-Mar-18	No fault	.
44	1	No fault	No fault	.	56	0	2019-Mar-18	No fault	.
45	0	2019-Mar-18	No fault	.	56	1	2019-Mar-18	2012-Nov-03	2326
45	1	2014-Mar-02	2013-Mar-06	361	57	0	2014-Aug-14	2012-Oct-26	657
45	1	2019-Mar-18	No fault	.	57	0	2019-Mar-18	2018-Apr-29	323
46	0	2015-Apr-01	2014-Jun-27	278	57	1	No fault	No fault	.
46	0	2017-Dec-01	No fault	.	58	0	2016-Dec-03	No fault	.
46	1	2013-Dec-24	2013-Mar-19	280	58	1	2014-Dec-27	No fault	.
46	1	2017-Dec-01	2014-Dec-07	1090	58	1	2016-Aug-13	2015-Oct-05	313
47	0	2019-Mar-04	2013-Nov-16	1934	59	0	2018-Nov-05	No fault	.
47	1	2017-Oct-22	No fault	.	59	1	No fault	No fault	.
48	0	2017-Jan-18	No fault	.	60	0	No fault	No fault	.
48	1	2015-Jun-19	2014-Apr-28	417	60	1	2014-Nov-19	No fault	.
48	1	2019-Mar-18	2017-Feb-25	751	60	1	2019-Mar-18	2016-May-31	1021
49	0	2017-Feb-10	No fault	.	61	0	2014-Sep-09	2013-Aug-02	403
49	0	2019-Mar-18	No fault	.	61	0	2015-Sep-25	2015-Jun-06	111
49	1	No fault	No fault	.	61	1	2014-Sep-09	No fault	.
50	0	2019-Mar-18	No fault	.	62	0	2017-Jun-24	2013-Nov-04	1328
50	1	No fault	No fault	.	62	0	2019-Mar-18	No fault	.
51	0	2014-Dec-13	2013-Sep-22	447	62	1	2014-Jun-06	2013-Dec-02	186
51	0	2019-Mar-18	No fault	.	62	1	2016-Sep-29	2016-Feb-22	220
51	1	2015-Mar-08	2013-Jul-10	606	63	0	2017-Nov-20	No fault	.
52	0	2015-May-28	2013-Sep-24	611	63	1	No fault	No fault	.
52	0	2017-Jul-14	No fault	.	64	0	2014-Sep-14	2014-Jun-21	85
52	1	2013-Apr-13	No fault	.	64	1	2013-Apr-27	2012-Dec-03	145
52	1	2013-Nov-04	No fault	.	64	1	2014-Sep-14	2014-Jun-18	88
52	1	2017-Mar-08	No fault	.	65	0	No fault	No fault	.
53	0	No fault	No fault	.	65	1	2019-Mar-18	No fault	.
53	1	2015-Sep-11	No fault	.	66	0	2015-Sep-25	2015-Apr-02	176
54	0	2017-Jan-04	No fault	.	66	1	2017-Jun-03	2015-Jul-05	699
54	1	2015-Sep-25	No fault	.					

[1] The date of change of the IFP module.
[2] Number of days the detection is made before the Ground Truth date.
[3] FDD Scheme in [8]

system performs on the same data used to test our scheme, showing that ALMA gains significantly in performance with this new method.

V. CONCLUSION

Harsh environments add significant stress to the operations of instruments and machinery. This effect is particularly true for the next-generation telescopes. They are generally located in places with extreme conditions that increase the personnel's risks, making operation and maintenance, particularly challenging. To reduce these risks and improve the operation's performance, developing fault detection schemes is crucial. These systems can help the personnel focus on high value-added tasks, reducing the possibility of malfunctions.

In this work, we developed a novel fault detection scheme for a slow degradation fault in the antenna array's communication pipeline at the ALMA radio telescope. To reduce false-positives due to the low signal-to-noise ratio, a noise reduction stage was developed and tested. We tuned the parameters of our filtering stage through GA and PSO training, improving the signal-to-noise ratio by more than twice.

In the fault detection stage, we took advantage of ESN's benefits to develop a detection scheme through prediction, time-shifting, and comparison to a fault-free lower bound. Our approach reduces the need to manually identify relevant features and signal dynamics to achieve fault detection. Moreover, the proposed approach does not require human intervention to identify relevant features or the signal dynamics in faulty and non-faulty modes to able to detect the fault later. This benefit reduces the requirements to the maintenance teams and could make the approach much more flexible in its application.

We tested our design with real offline monitoring data from the 132 IFPs available at the ALMA radio telescope. Although the signals' characteristics slightly varied among IFPs, our scheme was able to correctly detect the historical faults, with an accuracy of 70%. More importantly, the methodology only presented 3% of false negatives, which is extremely useful in this setting. Furthermore, the system was able to detect that the IFP was in a faulty mode more than a year earlier than the human operators were able to realize it.

Although not critical in this application, our scheme still has a high number of false-positives. Future work needs to be done to reduce this, mainly caused by the high level of noise in the raw data. Additionally, this scheme should be tested against data from other components to see if it can detect faults in different settings.

APPENDIX
SUMMARY CASES OF FAULT DETECTION

In this appendix, we present all the results from applying our fault detection scheme to the historical data of the IFPs. We used the data of all 66 antennas, each with two IFPs, with time-series ranging from 2012 to 2019.

The results of our approach are summarized in Tables 5 and 6. Both tables show, for each antenna and polarization under the ESN column, when the IFP's fault was detected by our scheme. In some cases, an IFP presented a fault more than once; hence it can appear multiple times.

Under the Ground Truth column, we present the date when the fault was identified by human operators, date that was validated by an engineering specialist at ALMA. Finally, the Δ column shows the difference, in days, between both dates.

Finally, Tables 7 and 8 give thorough details on how the current system, based on Kalman filters and expert knowledge, performs.

The Atacama Large Millimeter/submillimeter Array (ALMA), an international astronomy facility, is a partnership of the European Organisation for Astronomical Research in the Southern Hemisphere (ESO), the U.S. National Science Foundation (NSF) and the National Institutes of Natural Sciences (NINS) of Japan in cooperation with the Republic of Chile. ALMA is funded by ESO on behalf of its Member States, by NSF in cooperation with the National Research Council of Canada (NRC) and the National Science Council of Taiwan (NSC) and by NINS in cooperation with the Academia Sinica (AS) in Taiwan and the Korea Astronomy and Space Science Institute (KASI).

ALMA construction and operations are led by ESO on behalf of its Member States; by the National Radio Astronomy Observatory (NRAO), managed by Associated Universities, Inc. (AUI), on behalf of North America; and by the National Astronomical Observatory of Japan (NAOJ) on behalf of East Asia. The Joint ALMA Observatory (JAO) provides the unified leadership and management of the construction, commissioning and operation of ALMA.

REFERENCES

[1] R. Isermann, *Fault-Diagnosis Systems*. Berlin, Germany: Springer, 2006. [Online]. Available: http://link.springer.com/10.1007/3-540-30368-5

[2] R. A. Carrasco, F. Núñez, and A. Cipriano, "Fault detection and isolation in cooperative mobile robots using multilayer architecture and dynamic observers," *Robotica*, vol. 29, no. 4, pp. 555–562, Jul. 2011.

[3] V. T. Do and U.-P. Chong, "Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain," *Strojniski Vestnik—J. Mech. Eng.*, vol. 57, no. 9, pp. 655–666, Sep. 2011.

[4] F. Meinguet, P. Sandulescu, B. Aslan, L. Lu, N.-K. Nguyen, X. Kestelyn, and E. Semail, "A signal-based technique for fault detection and isolation of inverter faults in multi-phase drives," in *Proc. IEEE Int. Conf. Power Electron., Drives Energy Syst. (PEDES)*, Dec. 2012, pp. 1–6.

[5] Z. Germán-Salló and G. Strnad, "Signal processing methods in fault detection in manufacturing systems," *Procedia Manuf.*, vol. 22, pp. 613–620, Jan. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2351978918303858

[6] J. Duan, T. Shi, H. Zhou, J. Xuan, and Y. Zhang, "Multiband envelope spectra extraction for fault diagnosis of rolling element bearings," *Sensors*, vol. 18, no. 5, p. 1466, May 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/29738474

[7] H. Khorasgani, D. E. Jung, G. Biswas, E. Frisk, and M. Krysander, "Robust residual selection for fault detection," in *Proc. 53rd IEEE Conf. Decis. Control*, Dec. 2014, pp. 5764–5769.

[8] J. L. Ortiz and R. A. Carrasco, "Model-based fault detection and diagnosis in ALMA subsystems," *Proc. SPIE*, vol. 9910, Jul. 2016, Art. no. 99102S. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedingsof-spie/9910/2233204/Model-based-fault-detection-and-diagnosis-in-ALMA-subsystems/10.1117/12.2233204.full

[9] J. L. Ortiz and R. A. Carrasco, "ALMA engineering fault detection framework," *Proc. SPIE*, vol. 10704, Jul. 2018, Art. no. 107042K. [Online]. Available: https://www.spiedigitallibrary.org/conferenceproceedings-of-spie/10704/2312285/ALMA-engineering-fault-detectionframework/10.1117/12.2312285.full

[10] M. R. Napolitano, Y. An, and B. A. Seanor, "A fault tolerant flight control system for sensor and actuator failures using neural networks," *Aircr. Des.*, vol. 3, no. 2, pp. 103–128, Jun. 2000.

[11] L. Cork, R. Walker, and S. Dunn, "Fault detection, identification and accommodation techniques for unmanned airborne vehicles," in *Proc. Austral. Int. Aerosp. Congr. (AIAC)*, Melbourne, VIC, Australia, 2005, pp. 1–18. [Online]. Available: http://eprints.qut.edu.au/1729/1/1729.pdf

[12] M. A. Masrur, Z. Chen, B. Zhang, and Y. Lu Murphey, "Model-based fault diagnosis in electric drive inverters using artificial neural network," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Jun. 2007, pp. 1–7.

[13] A. Wootton, C. Day, and P. Haycock, "Echo state network applications in structural health monitoring," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Sep. 2014, pp. 1–7.

[14] S. Morando, M.-C. Marion-Péra, N. Y. Steiner, S. Jemei, D. Hissel, and L. Larger, "Fuel cells fault diagnosis under dynamic load profile using reservoir computing," in *Proc. IEEE Vehicle Power Propulsion Conf. (VPPC)*, Oct. 2016, pp. 1–6.

[15] Y. Fan, S. Nowaczyk, T. Rögnvaldsson, and E. A. Antonelo, "Predicting air compressor failures with echo state networks," in *Proc. 3rd Eur. Conf. Prognostics Health Manage. Soc. (PHME)*, I. Eballard and A. Bregon, Eds. Rochester, NY, USA: PHMSociety, 2016, pp. 568–578. [Online]. Available: http://hh.diva-portal.org/smash/get/diva2:948970/FULLTEXT01.pdf

[16] J. Westholm. (2018). *Event Detection and Predictive Maintenance Using Component Echo State Networks*. [Online]. Available: http://lup.lub.lu.se/student-papers/record/8931445

[17] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nature Commun.*, vol. 2, no. 1, p. 468, Sep. 2011, doi: 10.1038/ncomms1476.

[18] L. Appeltant, "Reservoir computing based on delay-dynamical systems," Ph.D. dissertation, Vrije Universiteit Brussel, Universitat de les Illes Balears, Instituto de Física Interdiscplinar y Sistemas Complejos IFISC (UIB-CSIC), Brussel, Belgium, May 2012. [Online]. Available: https://ifisc.uibcsic.es/users/phocus/attachments/AppeltantThesis_8mei_Hoofdletters.pdf

[19] A. Rodan and P. Tiňo, "Simple deterministically constructed cycle reservoirs with regular jumps," *Neural Comput.*, vol. 24, pp. 1822–1852, Mar. 2012.

[20] A. Czajkowski and K. Patan, "Robust fault detection by means of echo state neural network," in *Advanced and Intelligent Computations in Diagnosis and Control*, Z. Kowalczuk, Ed. Cham, Switzerland: Springer, 2016, pp. 341–352.
[21] A. Rodan and P. Tiňo, "Minimum complexity echo state network," *IEEE Trans. Neural Netw.*, vol. 22, no. 1, pp. 131–144, Jan. 2011.
[22] H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks-with an erratum note,"' German Nat. Res. Center Inf. Technol., Bonn, Germany, GMD Tech. Rep. 148, vol. 148, 2001. [Online]. Available: https://www.researchgate.net/publication/215385037_The_echo_state_approach_to_analysing_and_training_recurrent_neural_networks-with_an_erratum_note
[23] H. Jaeger, M. Lukoševičius, D. Popovici, and U. Siewert, "Optimization and applications of echo state networks with leaky-integrator neurons," *Neural Netw.*, vol. 20, no. 3, pp. 335–352, Apr. 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S089360800700041X
[24] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, Aug. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574013709000173
[25] M. Lukoševičius, "A practical guide to applying echo state networks," in *Neural Networks: Tricks of the Trade*. Berlin, Germany: Springer, 2012, pp. 659–686, doi: 10.1007/978-3-642-35289-8_36.
[26] J. Ortiz and J. Castillo, "Automating engineering verification in alma subsystems," *Proc. SPIE*, vol. 9149, Aug. 2014, Art. no. 914929.
[27] T. Dielman, "Choosing smoothing parameters for exponential smoothing: Minimizing sums of squared versus sums of absolute errors," *J. Mod. Appl. Stat. Methods*, vol. 5, pp. 117–128, May 2006.
[28] Z. Ismail and F. Y. Foo, "Genetic algorithm for parameter estimation in double exponential smoothing," *Austral. J. Basic Appl. Sci.*, vol. 5, no. 7, pp. 1174–1180, 2011.
[29] A. Chusyairi, R. N. S. Pelsri, and E. Handayani, "Optimization of exponential smoothing method using genetic algorithm to predict E-report service," in *Proc. 3rd Int. Conf. Inf. Technol., Inf. Syst. Electr. Eng. (ICITISEE)*, Nov. 2018, pp. 292–297.
[30] A. Simoni and E. Dhamo, "Evolutionary algorithm PSO and holt-winters method applied in hydro power plants optimization," in *Proc. Statist., Probab. Numer. Anal. Methods Appl. Conf.*, Albania, Dec. 2015. [Online]. Available: https://sites.google.com/a/fshn.edu.al/fshn/home/botim-special
[31] Y. Wang, H. Tang, T. Wen, and J. Ma, "A hybrid intelligent approach for constructing landslide displacement prediction intervals," *Appl. Soft Comput.*, vol. 81, Aug. 2019, Art. no. 105506. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1568494619302765
[32] M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998.
[33] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE ICNN*, vol. 4, Nov./Dec. 1995, pp. 1942–1948.

ANTHONY D. CHO received the B.S. degree in mathematics from the Universidad de Carabobo, Valencia, Venezuela, in 2008. He is currently pursuing the Ph.D. degree in industrial engineering and operations research with the Universidad Adolfo Ibáñez, Santiago, Chile.

His research interests include machine learning, evolutionary algorithms, operation research, prescriptive analytics, and image processing.

RODRIGO A. CARRASCO (Member, IEEE) received the degree in electrical engineering and the M.Sc. degree in engineering, focused in control systems, from the Universidad Católica de Chile, and the M.Phil. and Ph.D. degrees in industrial engineering and operations research from Columbia University.

He worked with Booz Allen Hamilton, leading operations research projects in Chile, Argentina, Brazil, Peru, and Canada. He was a Researcher with Siemens Corporate Research, Princeton, NJ, USA, developing decision support algorithms for smart grids and energy management. He is currently a Professor with the School of Engineering and Sciences, Universidad Adolfo Ibáñez, and the Academic Director of the Master in Industrial Engineering Program. He also founded and was the Initial Director of the UAI Systems Center, a center dedicated to technology transfer and solving complex real-life problems using operations research tools. His research is focused on the design and development of decision support tools and algorithms.

GONZALO A. RUZ (Member, IEEE) received the B.Sc., P.E., and M.Sc. degrees in electrical engineering from the Universidad de Chile, Santiago, Chile, in 2002 and 2003, respectively, and the Ph.D. degree from Cardiff University, U.K., in 2008. He is currently a Professor and the Research Director of the Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Santiago. His research interests include machine learning, evolutionary computation, data mining, gene regulatory network modeling, and complex systems.

JOSÉ LUIS ORTIZ received the B.Sc. and M.Sc. degrees in electrical engineering from the Universidad Católica de Chile, in 2003 and 2005, respectively. For the last ten years, he has worked with the ALMA Radio-Astronomical Observatory. He is currently a Senior Electronics Engineer with the Array Maintenance Group, where he has participated in the early stages of the telescopes assembly, integration, and verification effort, and now specializes in the design and implementation of techniques for array monitoring, remote troubleshooting, and advanced fault detection and diagnosis.