MDPI

Article

A Comprehensive Case Study of Data-Driven Methods for Robust Aircraft Sensor Fault Isolation

Nicholas Cartocci [1] , **Marcello R. Napolitano** [2] , **Gabriele Costante** [1] **and Mario L. Fravolini** [1,*]

1   Department of Engineering, University of Perugia, Via G. Duranti, 67, 06125 Perugia, Italy;
    nicholas.cartocci@unipg.it (N.C.); gabriele.costante@unipg.it (G.C.);

2   Department of Mechanical and Aerospace Engineering, West Virginia University,
    Morgantown, WV 26506-6106, USA; marcello.napolitano@mail.wvu.edu

*   Correspondence: mario.fravolini@unipg.it; Tel.: +39-393-21-62-358

Abstract: Recent catastrophic events in aviation have shown that current fault diagnosis schemes may not be enough to ensure a reliable and prompt sensor fault diagnosis. This paper describes a comparative analysis of consolidated data-driven sensor Fault Isolation (FI) and Fault Estimation (FE) techniques using flight data. Linear regression models, identified from data, are derived to build primary and transformed residuals. These residuals are then implemented to develop fault isolation schemes for 14 sensors of a semi-autonomous aircraft. Specifically, directional Mahalanobis distance-based and fault reconstruction-based techniques are compared in terms of their FI and FE performance. Then, a bank of Bayesian filters is proposed to compute, in flight, the fault belief for each sensor. Both the training and the validation of the schemes are performed using data from multiple flights. Artificial faults are injected into the fault-free sensor measurements to reproduce the occurrence of failures. A detailed evaluation of the techniques in terms of FI and FE performance is presented for failures on the air-data sensors, with special emphasis on the True Air Speed (TAS), Angle of Attack (AoA), and Angle of Sideslip (AoS) sensors.

Keywords: data-driven fault diagnosis; robust residual generation; fault isolation and estimation; Bayesian filtering; aircraft safety; flight data

1. Introduction

Fault Diagnosis (FD) is a term that describes the general problem of detecting and isolating faults in physical components or in the instrumentation of a system [1]. A broad class of FD methods is based on the concept of Analytical Redundancy (AR) [2]. The essence of AR methods lies in the comparison of the behavior of the system with the response of a mathematical model. Inconsistency between the actual and model response is considered a symptom of a possible failure. An FD scheme elaborates this information in real time to detect and isolate faults.

Mathematical models play a central role in AR-based FD; these are derived from either physical laws governing the dynamics of the system or are inferred directly from experimental data exploiting system identification techniques [3].

In the last 30 years, FD has been widely investigated, and many techniques have been proposed. Considering model-based approaches, the main research directions can be coarsely categorized as fault detection filters, parity-space schemes, diagnostic observers, and parameter estimation [1]. Classic methods dealing with model-based approaches can be found in [4–9]. More recently, approaches based on optimal criteria [10,11] and extensions to nonlinear [12], hybrid [13], and distributed systems [14] have been proposed. A collection of successful applications of model-based FD was presented in [15].

Data-based approaches are preferable instead in case the system dynamics is not precisely known and/or when the system input-output physical relations are too complex [16]. Currently, thanks to the advances in ICT technologies in both hardware and

software, substantial amounts of data are available for continuous monitoring and FD of engineering systems. This generated a great interest in FD techniques capable of handling large amounts of data. In this context, the most established and widespread techniques are Principal Component Analysis (PCA), Independent Component Analysis (ICA), Partial Least Squares (PLS), Fisher Discriminant Analysis (FDA), the Subspace-Aided approach (SAP), as well as their recent advancements [17]. Remaining in the data-based domain, FD diagnostic systems based on neural networks [18], computational intelligence [19], and machine learning [20,21] have attracted substantial interest in very recent years. Specifically, in the aviation and flight control systems community, there is abundant technical literature that clearly shows that the main reason for catastrophic accidents can be attributed, ultimately, to failures and/or malfunction of sensors, control surfaces, or components of the propulsion system, as reported in [22,23]. This explains the critical need for on-board FD systems able to promptly detect and identify the faulty components and, next, to enable a predetermined online failure recovery strategy [24]. Typically, hardware redundancy (featuring multiple sensors or actuators with the same function) and simple majority voting schemes are used to cope with FD. Although these methods are widespread, there is an increasing number of applications where the additional cost, weight, and size of the components are major constraints, such as small and autonomous flying vehicles. For these applications, AR-based FD methods are appealing approaches to increase flight safety. An excellent review on the state-of-the-art of AR-based FD methods in aerospace systems can be found in [25].

Regardless of the approach, the performance of an FD system is strictly related to the level of modeling uncertainty and measurement noise, whose presence implies that in fault-free conditions, the residual signals might significantly deviate statistically from zero. This entails that residuals must be robust with respect to uncertainties where the robustness of a residual is defined as the degree of sensitivity to faults compared to the sensitivity to uncertainty and noise [26]. Robustness issues become evident when dealing with experimental flight data [27–31]. Indeed, experimental data are affected by several additional sources of uncertainty such as, but not limited to, signal synchronization and quantization errors and imperfect positioning and calibration of the sensors [32,33]. Therefore, a fundamental step before the deployment of an FD scheme is its validation with actual flight data. Despite the relevance of these aspects in the mentioned studies, the problem of a detailed performance evaluation and comparison of Fault Isolation (FI) and Fault Estimation (FE) schemes using multi-flight data has been rarely fully addressed.

Building on the last considerations, the present study is entirely devoted to the setup and the quantitative comparison of well-known FI and FE methods using experimental sensor flight data. Specifically, different consolidated data-driven FI and FE schemes are applied to evaluate their performance on a set of 14 primary sensors of a semi-autonomous Tecnam P92 aircraft [34].

In the effort, experimental linear regression models of the sensors' response are identified from correlated measurements using multivariate linear regression techniques [3]. The identified models provide a natural set of primary residuals; however, the faults' directional properties are often not optimized to be used for sensor FI [35,36]. For this reason, enhanced residuals are derived by applying a linear transformation that allows better separation between the fault directions. The approach proposed in [37] that optimizes fault directions separation taking also into account experimental modeling errors is herein implemented.

Sensor FI is then performed using two different approaches. The first approach associates the fault with the sensor providing the smaller directional distance compared to the sensor fault directions [35]; the second approach, instead, is based on the reconstruction-based method proposed in [38].

A comprehensive (offline) analysis based on multi-flight data is performed to derive quantitative Fault Isolation (FI) and Fault Estimation (FE) performance as a function of the fault amplitudes.

Next, another study is performed to assess online sensor FI performance by monitoring the temporal history of the residuals and, using this information, to increase or decrease the belief in a fault hypothesis over time. For this purpose, a bank of recursive Bayesian filters is designed to infer in-flight sensors' fault probabilities. During online operation, a fault is declared when a fault probability reaches a defined threshold. This allows the computation of FI delays. The accumulated evidence in the probabilistic filters allows more reliable fault isolation, decreasing the false alarm rate at the expense of a small and acceptable delay in the diagnosis.

A detailed evaluation of offline and online FD performance is presented for failures on the air data sensors, including the True Air Speed (TAS, *TaS*), the Angle of Attack (AoA, *α*), and the Angle of Sideslip (AoS, *β*).

The present study is application oriented; therefore, the main contribution is not the development of a novel FD technique, but, rather, the design of consolidated data-driven sensor FD schemes and the quantitative evolution of fundamental quantities such as the fault isolation percentage, fault reconstruction accuracy, and in-flight fault isolation delays using multi-flight validation data. Therefore, the study provides a clear picture of the required design effort along with the achievable performance using consolidated FD schemes.

The paper is organized as follows. Section 2 introduces linear regression models for sensors' fault diagnosis; Section 3 introduces the FI and FE methods based on primary residuals, while Section 4 introduces the FI and FE methods based on transformed residuals. Section 5 introduces the Bayesian filters for online FI. Section 6 describes the multi-flight data, while Section 7 deals with the identification of experimental models. FI and FE results on validation data are discussed in Sections 8 and 9. Concluding and summary remarks are provided in Section 10.

2. Models for Sensor Fault Diagnosis

In this study, a set of n_x potentially faulty sensors is considered. The corresponding signals are concatenated in a vector $x(k) \in \mathbb{R}^{n_x}$; $u(k) \in \mathbb{R}^{n_u}$ is another vector of n_u signals functionally correlated with $x(k)$ such as control inputs, set-points, and other sensor measurements (assumed not to be faulty). It is assumed that in fault-free conditions, the sensor measurements $x(k)$ can be expressed in linear regression form as a function of $x(k)$ and $u(k)$, that is:

$$x_i(k) = \sum_{\substack{j=1 \\ j \neq i}}^{n_x} w_{xi,j} x_j(k) + \sum_{j=1}^{n_u} w_{ui,j} u_j(k) + \Delta_i(k) \quad i = 1...n_x \tag{1}$$

where $w_{xi,j}$ and $w_{ui,j}$ are the coefficients of the linear combination and $\Delta_i(k)$ characterizes modeling nonlinearity, uncertainty, and measurement noise concerning the *i*-th sensor. For simplicity, the linear models in (1) are rearranged as:

$$x_i(k) = w_{xi} x(k) + w_{ui} u(k) + \Delta_i(k) \quad i = 1...n_x \tag{2}$$

where $w_{xi} = [w_{xi,1}, ..., w_{xi,i-1}, 0, w_{xi,i+1}, ..., w_{xi,n_x}] \in \mathbb{R}^{n_x}$ and $w_{ui} = [w_{ui,1}, ..., w_{ui,n_u}] \in \mathbb{R}^{n_u}$. Putting the above n_x equations together, we get the following vector expression:

$$x(k) = \mathbf{W}_x x(k) + \mathbf{W}_u u(k) + \Delta(k) \tag{3}$$

where $\mathbf{W}_x = [w_{x1}; ...; w_{xn_x}] \in \mathbb{R}^{n_x \times n_x}$ and $\mathbf{W}_u = [w_{u1}; ...; w_{un_u}] \in \mathbb{R}^{n_x \times n_u}$ are constant matrices to be estimated from data (as discussed in Section 7). Model (3) has been widely used in the literature for estimating a sensor signal as a function of other correlated measurements, as shown, for instance, in [35,39–42]. The linear terms in (3) provide a linear estimation of the signals that are defined as:

$$\hat{x}(k) = \mathbf{W}_x x(k) + \mathbf{W}_u u(k). \tag{4}$$

The consistency of the measurements $x(k)$ are monitored through the vector of the primary residuals $r_0(k) \in \mathbb{R}^{n_x}$ that is defined as follows:

$$r_0(k) = x(k) - \hat{x}(k). \tag{5}$$

Substituting (4) into (5) leads to:

$$\begin{aligned} r_0(k) &= (\mathbf{I} - \mathbf{W}_x)x(k) - \mathbf{W}_u u(k) \\ &= \mathbf{W}x(k) - \mathbf{W}_u u(k) \end{aligned} \tag{6}$$

where $\mathbf{I} \in \mathbb{R}^{n_x \times n_x}$ is the identity matrix and $\mathbf{W} = \mathbf{I} - \mathbf{W}_x$. Using again (3) and (4) in (5), it is immediate to verify that:

$$r_0(k) = \Delta(k) \tag{7}$$

In other words, in fault-free conditions, the vector of the primary residuals is equal to the modeling error.

Sensor Fault Modeling

In this study, an additive (single) failure $f_i(k)$ on a generic (i-th) sensor of $x(k)$ is considered. Without any loss of generality, the failure can affect every single sensor. In the presence of a sensor failure, the vector $x(k)$ is replaced by its "faulty version", that is:

$$x(k) \leftarrow x(k) + \varepsilon_i f_i(k) \tag{8}$$

where $\varepsilon_i \in \mathbb{R}^{n_x}$ is the i-th column of the identity matrix $\mathbf{I}$. $f_i(k)$ is an arbitrary scalar, a function of time, modeling the fault shape; thus, $f_i(k)$ is zero before the fault occurrence and different from zero, starting from the "fault time instant" k_f. In this study, we consider a step fault:

$$f_i(k) = \begin{cases} 0 & t < k_f \\ A_i & t \geq k_f \end{cases} \quad i = 1...n_x \tag{9}$$

where A_i is the fault amplitude associated with the i-th sensor.

3. Sensor FI and FE Based on Primary Residuals

The first FI scheme is derived directly from the analysis of the primary residual vector $r_0(k)$ in (6) in the presence of a sensor fault. Substituting the faulty vector (8) in (6) leads to the faulty residual:

$$\begin{aligned} r(k) &= \mathbf{W}[x(k) + \varepsilon_i f_i(k)] - \mathbf{W}_u u(k) \\ &= r_0(k) + \mathbf{W}[\varepsilon_i f_i(k)]. \end{aligned} \tag{10}$$

Next, substituting in (10) the conditions (7) results in:

$$r(k) = \Delta(k) + w_i f_i(k) \tag{11}$$

where the vector $w_i \in \mathbb{R}^{n_x}$ is the i-th *column* of the matrix $\mathbf{W}$. Vector w_i is known as the "fault signature" or the "fault direction" associated with the sensor fault $f_i(k)$. In the theoretical case of zero modeling error in (11), this results in:

$$r(k) = w_i f_i(k). \tag{12}$$

In other words, the vector $r(k)$ is exactly parallel to vector w_i. This directional information can be used for sensor fault isolation purposes.

3.1. Mahalanobis Distance-Based FI and FE

This technique exploits the directional property of the n_x primary residuals in (12) to formulate a sensor FI hypothesis [43]. Specifically, in the presence of a fault $f_i(k)$, Equation (12) implies that the residual vector $r(k)$ assumes a specific known direction in the residual space $\mathbb{R}^{n_x}$ given by the i-th column vector w_i of matrix $\mathbf{W}$. By taking

advantage of this fault directional property, it is possible to identify the faulty sensor by comparing the direction of the current residual vector $r(k)$ with the known n_x directions w_i and assigning the fault to the sensor whose direction is the closest to the $r(k)$ direction. In practice, this FI logic is implemented as follows. As the first step, in order to be independent of the fault amplitude, both the residual $r(k)$ and the sensor fault directions are normalized to a unity norm. Specifically, at each instant k, the normalized residual $(||\bar{r}(k)||_2 = 1)$ is computed as $\bar{r}(k) = r(k)/[(r^T(k)r(k))]^{0.5}$, while the normalized fault directions $(||\bar{w}_i(k)||_2 = 1 \quad i = 1...n_x)$ are defined as $\bar{w}_i = w_i/[(w_i^T w_i)]^{0.5}$. Next, since the faults can have either a positive or a negative amplitude, it was deemed necessary to introduce negative normalized fault directions $(-\bar{w}_i \ i = 1...n_x)$, implying that the overall number of fault directions is $2n_x$. The distance between the normalized residual $\bar{r}(k)$ and the normalized fault directions $\bar{w}_i$ can be measured using different norms such as the Euclidean distance, the angular distance, or the Mahalanobis distance [44,45]. In this study, we selected the Mahalanobis distance since this approach allows accounting for modeling error information derived from experimental data. The error matrix $\mathbf{G}(k) \in \mathbb{R}^{n_x \times 2n_x}$ of the differences between the normalized residual $\bar{r}(k)$ and the normalized fault directions is defined as:

$$\mathbf{G}(k) = [\bar{r}(k) - \bar{w}_1, ..., \bar{r}(k) - \bar{w}_{n_x}, \ \bar{r}(k) + \bar{w}_1, ..., \bar{r}(k) + \bar{w}_{n_x}] \tag{13}$$

where the first n_x columns are associated with positive faults and the last n_x with negative faults. Next, the Mahalanobis distance associated with the columns $g_j \in \mathbb{R}^{n_x}$ of $\mathbf{G}(k)$ is defined as:

$$e_j(k) = \sqrt{g_j^T \mathbf{S}_r^{-1} g_j} \quad j = 1...2n_x \tag{14}$$

where $\mathbf{S}_r \in \mathbb{R}^{n_x \times n_x}$ is the covariance matrix of the residuals (for its computation, see Note 1). At each sample time instant k (following a fault detection), the FI is performed associating the fault with the sensor leading to the minimum value of the distances in (14). That is, defining $j_m(k)$ as:

$$j_m(k) = \underset{j=1...2n_x}{argmin}(e_j(k)) \tag{15}$$

then the faulty sensor index is identified as follows:

$$i_F(k) = \begin{cases} j_m(k) & if \quad 1 \leq j_m(k) \leq n_x \\ j_m(k) - n_x & if \quad n_x + 1 \leq j_m(k) \leq 2n_x \end{cases}. \tag{16}$$

The amplitude of the fault is positive for $1 \leq j_m(k) \leq n_x$, and it is negative for $n_x + 1 \leq j_m(k) \leq 2n_x$. Once a fault has been associated with the sensor $i_F(k)$, its amplitude can be directly estimated considering the i_F-th component $r_{iF}(k)$ of the primary residual vector $r(k)$. In fact, from (11), this results in:

$$r_{iF}(k) = \Delta_{iF}(k) + w_{(iF,iF)} f_{iF}(k). \tag{17}$$

Next, since by construction, $w_{(iF,iF)} = 1 \ \forall i_F(k)$, this results in:

$$r_{iF}(k) = \Delta_{iF}(k) + f_{iF}(k). \tag{18}$$

In other words, $r_{iF}(k)$ provides a direct (noisy) estimation of the fault amplitude $\hat{A}_i(k) = \hat{f}_{iF}(k)$ defined in (9), that is:

$$r_{iF}(k) = \hat{f}_{iF}(k) = \hat{A}_i(k). \tag{19}$$

3.2. Reconstruction-Based FI and FE

Another approach for FI and FE is the well-known reconstruction-based method [38,46,47]. In this approach, for each fault direction w_i, the fault amplitude, at time

instant k, is estimated by computing the fault amplitude $f_i(k)$ that minimizes the norm-two of the reconstruction errors defined as:

$$e_i(k) = [(r(k) - w_i f_i(k))^T (r(k) - w_i f_i(k))]^{0.5} \quad i = 1...n_x. \tag{20}$$

The optimal values for $f_i(k)$ are derived by computing the minimum of the quadratic form (20) with respect to the scalar $f_i(k)$. This leads to the closed-form solution:

$$\hat{f}_i(k) = \frac{w_i^T r(k)}{w_i^T w_i} \quad i = 1...n_x. \tag{21}$$

It is observed that the fault amplitude estimation $\hat{f}_i(k)$ provided by (21) can be either positive or negative, implying that the reconstruction-based approach handles both positive and negative faults. The faulty sensor is then isolated associating the fault with the sensor leading to the minimum of the "reconstructed" errors achieved substituting the fault estimates (21) in (20). Therefore, the sensor fault index is:

$$i_F(k) = \underset{j=1...n_x}{argmin}(e_j(k)). \tag{22}$$

Similar to (19), the estimated fault amplitude is:

$$\hat{A}_i(k) = r_{iF}(k). \tag{23}$$

Note 1: The residual noise covariance matrix $\mathbf{S}_r$ in (14) depends on the modeling uncertainty $\Delta(k)$ in (7), which is very difficult to characterize "a priori". For this reason, $\mathbf{S}_r$ was estimated from the experimental data. From (7), it is inferred that in fault-free conditions, the relationship $r_0(k) = \Delta(k)$ holds. This allows estimating $\mathbf{S}_r$ as the sample covariance matrix of the residual vector inferred from experimental flight data in fault-free conditions. The experimental $\mathbf{S}_r$ matrix is therefore:

$$\hat{\mathbf{S}}_r = \text{cov}\{r_0(k)\} = \text{cov}\{\Delta(k)\} = \hat{\mathbf{S}}_\Delta \quad k \in \text{train\_data}. \tag{24}$$

Note 2: In this effort, we assumed that faults originate from sensors in an additive fashion. Under this hypothesis, it is possible to isolate the faulty sensor by analyzing the fault direction of the residual signals as shown in Equation (9). It is worth emphasizing that primary residuals can also be affected by the occurrence of generic internal system faults that in turn produce a specific fault signature on the residuals. In the case of internal faults, it is necessary to reconstruct the roots of the fault symptoms in the subsystems that constitute a complex system in order to be able to isolate the faulty component. In the event an internal fault produces a signature identical to those of a sensor fault, they would not be distinguishable by our scheme. The isolation of internal faults requires the knowledge of the internal dynamics of the subsystems and its cause-effect relationships. This problem, although interesting, is beyond the scope of this study, which is limited to addressing the sensor fault isolation problem. An interesting review on the cause and fault propagation can be found in [48].

4. Sensor FI and FE Based on Transformed Residuals

Primary residuals are not optimized from an FI point of view. For this reason, they are usually processed by applying transformations to achieve suitable new fault directions that facilitate the FI. An extensive literature on transformed directional residuals techniques is available. One of the first applications is the Beard–Jones filter [13], essentially a Luenberger observer whose gains are selected so that the directions of the residuals can be advantageously used to identify faulty sensors. In [49], taking advantage of the properties of un-observability subspaces, a set of residual transformations that are unaffected by all faults except one was proposed. In this context, the methodology outlined in [36] is also

relevant, where the interaction between directional residuals and fault isolation properties was analyzed.

The FI methodology proposed in [37] was considered in this effort because it provides optimized robust performance considering the directional properties of the residual noise covariance matrix $\mathbf{S}_r$. This feature is particularly important when dealing with experimental noisy data. In the approach in [37], a linear transformation of the primary residual vector $r(k)$ is introduced to provide optimized performance with respect to noise immunity. A transformed residual $\rho(k) \in \mathbb{R}^{n_x}$ with the same number of elements of $r(k)$ is defined as:

$$\rho(k) = \mathbf{W_t} r(k) = \mathbf{W_t}[\mathbf{W}x(k) - \mathbf{W}_u u(k)] \tag{25}$$

where $\mathbf{W_t} \in \mathbb{R}^{n_x \times n_x}$ is the transformation matrix to be computed. Considering (25), the new fault directions associated with $\rho(k)$ are the columns w_{ρ_i} of the matrix:

$$\mathbf{W}_\rho = \mathbf{W_t}\mathbf{W} \quad \in \mathbb{R}^{n_x \times n_x}. \tag{26}$$

The covariance of $\rho(k)$ is by definition:

$$\mathbf{S}_\rho = \text{cov}\{\mathbf{W_t}\, r(k)\} = \mathbf{W_t}\, \mathbf{S}_r\, \mathbf{W_t}^{\mathbf{T}} \tag{27}$$

where the matrix $\mathbf{W_t}$ is computed applying the method proposed in [37]. Considering this approach, let $\mathbf{A} \in \mathbb{R}^{n_x \times n_x}$ and $\mathbf{B} \in \mathbb{R}^{n_x \times n_x}$ be two symmetric matrices. Assuming $\mathbf{A}$ positive definite, it can be shown that there exists a matrix $\mathbf{V} \in \mathbb{R}^{n_x \times n_x}$ such that:

$$\mathbf{V^T A\, V = I}, \tag{28}$$

$$\mathbf{V^T\, B\, V} = \mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n) \tag{29}$$

where $\mathbf{V}$ and $\mathbf{\Lambda}$ are the solutions of the generalized eigenvalue problem. Specifically, the columns of $\mathbf{V}$ are the eigenvectors of the matrix $\mathbf{A}^{-1}\mathbf{B}$, while the columns of $\mathbf{\Lambda}$ are the corresponding eigenvalues. In the present study, setting $\mathbf{A} = \mathbf{S}_r$, $\mathbf{B} = \mathbf{W}\mathbf{W}^{\mathbf{T}}$, and $\mathbf{V}^{\mathbf{T}} = \mathbf{W_t}$, Equations (28) and (29) become, respectively:

$$\mathbf{W_t}\, \mathbf{S}_r\, \mathbf{W_t}^{\mathbf{T}} = \mathbf{I}, \tag{30}$$

$$\mathbf{W_t}\mathbf{W}\mathbf{W^T}\mathbf{W_t}^{\mathbf{T}} = \mathbf{W}_\rho \mathbf{W}_\rho{}^T = \mathbf{\Lambda}. \tag{31}$$

Relationship (30) implies that the noise covariance matrix of the transformed residuals in (27) is spherical, that is $\mathbf{S}_\rho = \mathbf{I}$. This property is critical since in the presence of a spherical symmetry noise, the optimal decision line between two fault directions is simply the bisector [50]. Thus, this property will be exploited for the design of the FI algorithm in Section 4.1.

As explained in Note 1, the $\mathbf{S}_r$ matrix is estimated using the experimental $\hat{\mathbf{S}}_r$ in (24), and the generalized eigenvalues $\mathbf{\Lambda}$ are derived using commercially available scientific software.

4.1. Transformed Residuals Based on FI and FE

Fault isolation based on the transformed residual $\rho(k)$ is performed using the same approaches used for the primary residuals in Sections 3.1 and 3.2. Specifically, defining the new transformed residual directions w_{ρ_i} $i = 1...n_x$ as the columns of the matrix $\mathbf{W}_\rho$, the transformed error matrix $\mathbf{G}_\rho(k) \in \mathbb{R}^{n_x \times 2n_x}$ is defined as:

$$\mathbf{G}_\rho(k) = [\bar{\rho}(k) - \bar{w}_{\rho_1}, ..., \bar{\rho}(k) - \bar{w}_{\rho_{n_x}}, \bar{\rho}(k) + \bar{w}_{\rho_1}, ..., \bar{\rho}(k) + \bar{w}_{\rho_{n_x}}] \tag{32}$$

where $\bar{\rho}(k)$ is the normalized transformed residual derived from $\rho(k)$ and $\bar{w}_{\rho_i}$ $i = 1...n_x$ are the new normalized transformed fault directions derived from w_{ρ_i}. The distance errors associated with the columns g_{ρ_i} of the matrix $\mathbf{G}_\rho(k)$ are defined as:

$$e_{\rho_j}(k) = \sqrt{g_{\rho_j}{}^T g_{\rho_j}} \quad j = 1...2n_x. \tag{33}$$

The fault isolation index is determined by applying the same technique introduced in Section 3.1, that is:

$$j_m(k) = \underset{j=1...2n_x}{argmin}\left(e_{\rho_j}(k)\right). \tag{34}$$

It is observed that the decision method (34) based on the Euclidean distances (33) is equivalent to treating as the decision line the bisectors between the directions g_{ρ_i} in (32), which, in the case of spherical noise ($\mathbf{S}_r = \mathbf{I}$), are also optimal for testing fault isolation [50].

The faulty sensor $i_F(k)$ is isolated applying (16) to $j_m(k)$ derived from (34). Finally, the fault amplitude is estimated using again (19), that is:

$$r_{iF}(k) = \hat{A}_i(k) = \hat{f}_{iF}(k). \tag{35}$$

Note 3: In [37], it was shown that the best performances are achieved in the case that the transformation matrix $\mathbf{W_t}$ leads to the diagonalization, not only of the noise covariance matrix $\mathbf{S}_\rho$, but also of the transformed fault direction matrix $\mathbf{W}_\rho$. This may be possible only if the number of residuals is larger than the number of sensors. Alternatively, approximated robust FI methods based on optimality concepts as those in [35,51] can be applied.

4.2. Transformed Residuals with Reconstruction-Based FI and FE

Consistent with the approach of Section 3.2, the reconstruction error for transformed residuals is defined as:

$$e_{\rho_i}(k) = [(\rho(k) - w_{\rho_i} f_i(k))^T (\rho(k) - w_{\rho_i} f_i(k))]^{0.5} \quad i = 1...n_x \tag{36}$$

while the reconstructed fault amplitude is:

$$\hat{f}_i(k) = \frac{w_{\rho_i}^T \rho(k)}{w_{\rho_i}^T w_{\rho_i}} \quad i = 1...n_x. \tag{37}$$

Fault isolation is again performed by computing the minimum of the errors in (36); in other words, the fault isolation index is:

$$i_F(k) = \underset{j=1...n_x}{argmin}\left(e_{\rho_j}(k)\right). \tag{38}$$

Finally, similar to (23), the estimated fault is:

$$\hat{A}_i(k) = \hat{f}_{iF}(k). \tag{39}$$

5. Bayesian Filtering for Online Fault Isolation

The FI approaches described in Sections 3 and 4 are based on a decision method that isolates the faulty sensor as the one providing the minimum of a specific error measure inferred from the residual. This FI logic is based only on information at the current sample time k and does not take into account the history of the residuals in the previous instants. On the other side, this information is useful for increasing or decreasing the belief of the FI decision over time.

In this paper, we propose a Bayesian Filter (BF) approach for managing the stream of information coming from sensors' measurements. Specifically, we implemented the so-called "discrete Bayes filter" proposed in [52]. This filter is essentially a recursive algorithm

used to estimate the distribution of a discrete probability function. This type of filter was selected because it builds on a very solid and comprehensible theoretical background; an introduction to the Bayesian decision theory can be found in [53]. In the present study, the estimated distribution models the probability (belief) that a generic sensor is faulty. The BF infers sensor fault probabilities by processing recursively the error information $e_i(k)$ $i = 1...n_x$ defined in (14), (20), (33) and (36) for the different methods.

Two possible operational statuses are assumed for the sensors. State F_i indicates the events for which the i-th sensor is faulty, while state N_i the event for which it is not faulty. Let $p(e_i(k)|F_i(k))$ be the likelihood function representing the probability of observing an error $e_i(k)$ given that the i-th sensor is faulty; a similar definition holds for $p(e_i(k)|N_i(k))$. According to Bayes' theorem, the posterior fault probabilities $p(F_i(k)|e_i(k))$ are given by:

$$p(F_i(k)|e_i(k)) = \frac{p(e_i(k)|F_i(k))p(F_i(k))}{p(e_i(k)|F_i(k))p(F_i(k)) + p(e_i(k)|N_i(k))p(N_i(k))} \qquad i = 1...n_x \quad (40)$$

$$p(N_i(k)|e_i(k)) = 1 - p(F_i(k)|e_i(k))$$

where $p(F_i(k))$ and $p(N_i(k))$ are the a priori fault and non-fault probabilities, respectively, for sensor i. A key step in the above inference mechanism is the definition of the likelihood functions $p(e_i(k)|F_i(k))$. These have to be designed to model the probability of experiencing a distance error $e_i(k)$ given that the i-th sensor is faulty; therefore, their distribution needs to be maximum for $e_i(k) = 0$ and should decrease as $e_i(k)$ increases. Similar reasoning applies to $p(e_i(k)|N_i(k))$. This behavior is captured by the following likelihood functions [54–57]:

$$p(e_i(k)|F_i(k)) = e^{-\alpha_i e_i(k)},$$
$$p(e_i(k)|N_i(k)) = e^{-\alpha_i [e_i(k)]^{-1}} \qquad (41)$$

where the parameters α_i are used to regulate the shape of the probability functions. At each time step, following the computation of the posterior probabilities (40), the a priori probabilities are updated recursively using:

$$p(F_i(k+1)) = p(F_i(k)|e_i(k)),$$
$$p(N_i(k+1)) = p(N_i(k)|e_i(k)) \qquad (42)$$

to propagate the current probability information to the next time step. In this study, the probabilistic filters are activated immediately after a fault is detected at $k = k_f$. At instant k_f, it is assumed that all the n_x sensors have the identical probability to be faulty; in other words, the filters are initialized with:

$$p\left(F_i(k_f)\right) = 1/n_x$$
$$p\left(N_i(k_f)\right) = 1 - p\left(F_i(k_f)\right) \qquad i = 1...n_x. \quad (43)$$

5.1. Probabilistic (Online) Fault Isolation Method

The n_x recursive BF probabilities are employed for online FI, which is associated with in-flight conditions. Following a fault detection at $k = k_f$, the fault is assigned to the sensor whose fault probability $p(F_i(k)|e_i(k))$ first reaches a defined threshold. If $p(F_i(k)|e_i(k))$ exceeds the threshold, the FI scheme has "good confidence" that a fault has occurred on the i-th sensor. Therefore, a failure on that sensor is declared. In this study, the threshold was empirically set at 0.7 (70%); see Note 4.

5.2. Probability Function Tuning

The performance of the BFs depends strictly on the shape of the probability functions in (42), which, in turn, depend on the values of the parameters α_i. These values have to be carefully selected to make the filters sensitive to faults while limiting false alarms. Interestingly, a significant difference in the tuning of the filters was found for

the distance-based and reconstruction-based techniques. This difference is essential because, while the first method operates with normalized residuals, the second operates with non-normalized residuals.

5.2.1. Distance-Based Methods' Tuning

For the distance-based techniques, it was relatively simple to calibrate α_i to limit false alarms. These techniques being based on normalized residuals ($|\bar{r}(k)| = 1$, $|\bar{\rho}(k)| = 1$), the α_i values were inferred from the mean value (in fault-free conditions) of the distance between the normalized residual and normalized fault directions. Then, for distances $e_i(k)$ smaller than the mean value of the distance, the α_i values of $p(F_i(k)|e_i(k))$ in (41) are tuned to be larger than 0.5, so that the BF increases the fault belief with respect to the previous step. Conversely, for larger distances, $p(F_i(k)|e_i(k))$ is lower than 0.5, resulting in a decrease in the belief. This tuning method was found to be effective since the distance $e_i(k)$ (independent of the fault size) is always contained in the hyper-ellipsoid (14).

5.2.2. Reconstruction-Based Methods' Tuning

For reconstruction-based techniques, the application of the above calibration technique is not possible. This is due to the fact that the errors $e_i(k)$ in (20) derive from non-normalized residuals ($r(k)$ and $\rho(k)$), and therefore, their values are proportional to the fault amplitude. Thus, the selection of α_i values that are suitable for any fault amplitude is not trivial. To overcome the problem, we opted for a "forced" normalization of the errors $e_i(k)$ in (20) to be used as the input of the BFs. The following normalization method was applied. At each sample instant k, the 14 errors $e_i(k)$ are ranked so that $e_{rank_1}(k) \leq e_{rank_2}(k) \leq e_{rank_3}(k)... \leq e_{rank_{14}}(k)$; next, normalized errors $\bar{e}_i(k)$ are defined as $\bar{e}_i(k) = e_i(k)/e_{rank_2}(k)$. Through this scheme, the new $\bar{e}_i(k)$ are such that one normalized error is always less than one; another is exactly equal to one, while all the remaining normalized errors are larger than one. Using the normalized $\bar{e}_i(k)$, it is possible to define a simple tuning strategy for α_i. Specifically, α_i are tuned such that the fault belief increases for the (nearest) sensor having $\bar{e}_i(k) \leq 1$, remains unchanged for the sensor having $\bar{e}_i(k) = 1$, and decreases for all the remaining sensors. It is immediate to verify that $\alpha_i = \ln 2$ guarantees this behavior. The normalized errors $\bar{e}_i(k)$ are then used in the BFs (40)–(43) (in place of $e_i(k)$) for the reconstruction based methods.

Note 4: An important aspect in the design of a fault diagnosis scheme is the definition of the threshold values to be used for fault detection and isolation. Clearly, the values of these thresholds have a direct impact on the missed alarm rate and false alarm rate. Although several methods have been introduced and tested to compute optimized thresholds [58,59], most of the detection methods require a priori knowledge of the signal distribution, changed parameters, and the change amplitude (CUSUM, likelihood ratio test, etc.). Furthermore, these methods assume that modeling errors are Independent and Identically Distributed (IID) random variables. Unfortunately, in our study, we found that the above-mentioned assumptions are not satisfied by experimental residuals and that the application of theoretical thresholds produces very conservative results lacking any practical utility. For this reason, these values were set empirically by trial and error.

Note 5: Since FI algorithms operates in real time, important factors are their computational and memory space requirements. Given that the transformation of fault signatures is performed offline, techniques based on transformed and primary residuals require exactly the same memory space for the storage of the models. As for the computational cost, in reconstruction-based techniques, the computational complexity is $O(n^2)$, and the memory space required is $3n_x$ (n_x to store residuals, n_x to store estimated fault amplitudes, and n_x to store errors). In distance-based techniques, the computational complexity, instead, depends on the operations requested to compute the errors in (20) or (36). For the technique with primary residuals, the complexity is $O(n^3)$ because the matrix product involving Sr^{-1} in (14), while in the case of transformed residuals, it is equal to $O(n^2)$. The memory space required is quantifiable in $6n_x$ (n_x to store the residuals, n_x to store the normalized

residuals, $2n_x$ to store $G(k)$ (or $G_\rho(k)$), and $2n_x$ to store the errors). The overall results are summarized in Table 1. The Bayesian filtering introduces a computational complexity $O(n)$ to calculate the posterior fault probabilities and requires a memory space $4n_x$ ($2n_x$ to store the likelihood probabilities and $2n_x$ to store the posterior fault probabilities).

Table 1. Computational complexity and memory space required by the techniques.

	Maha. Pr.	**Maha. Tr.**	**RBEPr.**	**RBE Tr.**
Computational complexity	$O(n^3)$	$O(n^2)$	$O(n^2)$	$O(n^2)$
Memory space	$6n_x$	$6n_x$	$3n_x$	$3n_x$

6. Aircraft and Flight Data

The considered FI and FE techniques were designed and validated using sensor flight data from a Tecnam P92 aircraft, shown in Figure 1. The aircraft mass is approximately 600 kg, and the propulsion is provided by a 74 kW Rotax 912 ULS with a two-blade fixed-pitch propeller, for a maximum cruise speed of 219 km/h and an operational ceiling of 4200 m. Data were acquired in semi-autonomous mode; in other words, the aircraft was manually flown by a pilot during the take-off and landing phase, while flown autonomously in cruise flight conditions. A set of nine flight datasets was considered in this study, of which five flights (with a total length of 2 h and 20 min) were used for the design of the FI and FE schemes, while the remaining four (with a total length of 2 h and 5 min) were used for validation purposes. These proprietary flight data previously acquired as part of an internal industrial research study for aircraft certification purposes were provided courtesy of Tecnam corporation. The data sampling time was set at 0.1 s. The study did not consider data associated with take-off, initial climb, final descent, and landing phases for the specific reason that in those flight conditions, the aerodynamic behavior of the aircraft is quite different with respect to cruise conditions due to the deployment of flaps. The use of multi-flight data for the training and the validation tasks is extremely important because, as for all data-driven techniques, the reliability of the results strongly depends on the completeness of the data, their spectral richness, and the coverage of all operating conditions. A total of 20 sensors was considered in this study, and they are listed in Table 2. The FI schemes were designed for the set x_0 of the 14 sensors indicated in Table 2 (the $x(k)$ signals in (3)), while the six actuation signals u_0 were the control deflections and engine commands (the $u(k)$ signals in (3)).

Table 2. Aircraft sensors.

	x_0				u_0
α	Angle of attack	P	Roll speed	Ap	Aileron position
β	Drifting angle	Q	Pitch speed	Rp	Rudder position
TaS	True air speed	R	Yaw speed	Tp	Thrust lever position
ϕ	Roll angle	NNx	Longitudinal load factor	Pp	Pitch trim position
θ	Pitch angle	NNy	Lateral load factor	Sp	Stabilator position
ψ	Yaw angle	NNz	Vertical load factor	Eng	Engine revolution
$ZLaser$	Altitude laser	$Zpos$	GPS altitude		

Data Normalization

The experimental data listed in Table 2 have quite different ranges and orders of magnitude; this issue suggested using a data normalization to zero mean and unit standard deviation.

The presence of regressor signals with very different operative ranges could cause the exclusion of the informative regressors with lower power from the model. This problem is

avoided through normalization, which makes all signals with zero mean and unit variance. For each a signal, the normalization was performed using:

$$z_i(k) = \frac{z_{0_i}(k) - \mu_{z_{0_i}}}{\sigma_{z_{0_i}}} \quad i = 1 \ldots (n_x + n_u) \tag{44}$$

where $z_{0_i}(k)$ is the *i*-th non-normalized signal ($x_{0_i}(k)$ or $u_{0_i}(k)$); $\mu_{z_{0_i}}(k)$ is its mean, and $\sigma_{z_{0_i}}$ is its standard deviation. Since the $z_i(k)$ signals are normalized, this implies that faults in (8) are associated with normalized signals; therefore, the fault amplitude estimations provided by (19), (23), (37) and (39) must be de-normalized to recover the actual fault amplitude. This can be achieved inverting (44), so that, assuming $\hat{f}_i(k)$ to be the normalized fault reconstructed through (19), (23), (37) and (39), the actual amplitude of the fault is:

$$\hat{f}_{i_{den}}(k) = \sigma_{x_{0_i}} \hat{f}_i(k). \tag{45}$$

Figure 1. The semi-autonomous Tecnam P92 aircraft.

7. Experimental Models for Sensor FI

The fault isolation techniques introduced in Sections 3 and 4 require the definition of linear multivariate models. Specifically, the definition of the matrices $\mathbf{W}_x$ and $\mathbf{W}_u$ associated with the $n_x = 14$ linear models in (4) is required. For each one of the n_x sensors, the model was identified separately using standard system identification techniques [3]. From (3), the resulting models are:

$$\hat{x}_i(k) = w_{x_i} x(k) + w_{u_i} u(k) \quad i = 1 \ldots 14 \tag{46}$$

where $w_{xi} \in \mathbb{R}^{14}$ and $w_{ui} \in \mathbb{R}^6$ are the *i*-th row of the matrices $\mathbf{W}_x$ and $\mathbf{W}_u$, respectively. Each of these 14 models depends on 20 potential regressors. It is common practice in model identification to select a subset of regressors that are critical for the estimation of the predicted output $\hat{x}_i(k)$. The regressors' selection step was partially automated using the algorithm known as the "stepwise regressor selection method". This is a well-known iterative data-driven algorithm for implementing a linear model by successively adding and/or removing regressors based on their statistical significance in a regression model [60]. In this

effort, the stepwise selection was based on the training data using the ad-hoc procedure available in [61]. For each model, the selection of the best set of regressors was formulated through the analysis of the Root Mean Squared Error (RMSE) of the prediction error defined as:

$$RMSE_i = \sqrt{mean\left[\left(x_i(k) - \hat{x}_i(k)\right)^2\right]} \quad i = 1...14. \tag{47}$$

Figure 2 shows the evolution of the RMSE for the models of the sensors α, β, and *TaS* evaluated both on the training and the validation flight data as a function of the number of regressors included in the model. As expected, while the training data result in a monotone decrease of the RMSE with the increase of the number of regressors in the model, for the validation data, the RMSE reaches a minimum; also, a larger number of regressors induces a decrease of the prediction accuracy. This is a typical example of the well-known modeling overfitting problem. Consequently, for each sensor, the best model was identified as the one that produced the minimum RMSE on the validation data set. The sets of regressors selected by this approach for the α, β, and *TaS* models are reported in Table 3. Without any loss of generality, this procedure was also applied to the other remaining 11 sensors. The number of regressors selected for the 14 models is reported in Table 4.
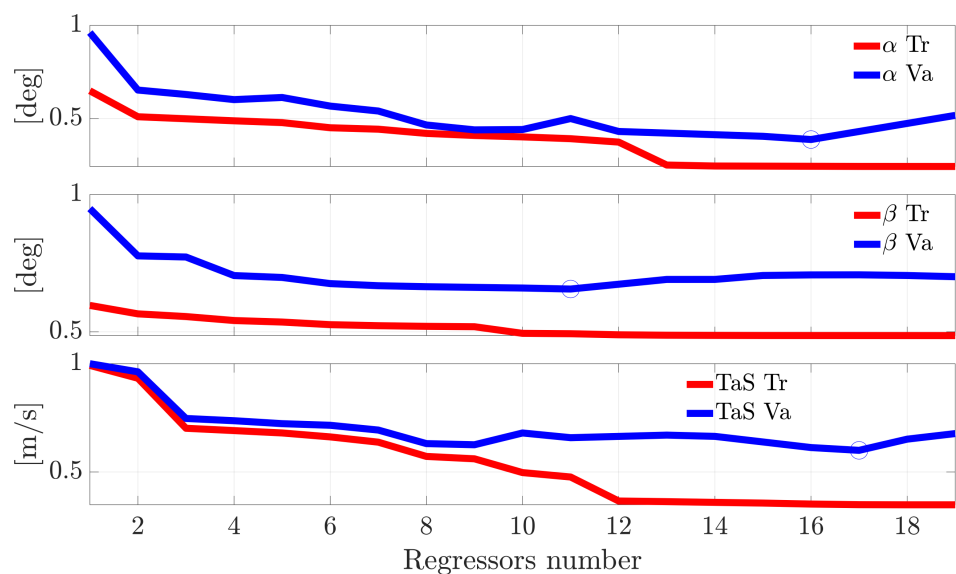
Figure 2. RMSE as function of the number of regressors in the model for the $\alpha(k)$, $\beta(k)$, and *TaS*(k) sensors. Training (Tr) and Validation (Va) data.

Table 3. Selected regressors by the stepwise method for the α, β, and *TaS* models. TAS, True Air Speed.

Sensor																
							Selected Regressors (by the Stepwise Method)									
α	β	TaS	NNx	NNy	NNz	P	Q	R	ZLaser	Z	Ap	Rp	Tp	Pp	Sp	Eng
β	α	TaS	NNx	NNy	NNz	P	R	ϕ	ZLaser	Ap	Tp					
TaS	α	β	NNx	NNy	NNz	P	Q	R	ϕ	ψ	Zpos	ZLaser	Ap	Tp	Pp	Sp	Eng

Table 4. Number of selected regressors by the stepwise method for the 14 prediction models.

α	β	**TAS**	**NNx**	**NNy**	**NNz**	**P**	**Q**	**R**	ϕ	θ	ψ	**Zpos**	**ZLaser**
16	11	17	16	12	18	14	16	8	9	8	11	15	10

Figure 3 compares the RMSE for the 14 models produced by the (normalized) training and validation flight data. It is observed, for all models, that the training and the validation

performance are comparable; this implies that the adopted regressor selection procedure is substantially correct and successfully avoids overfitting problems.
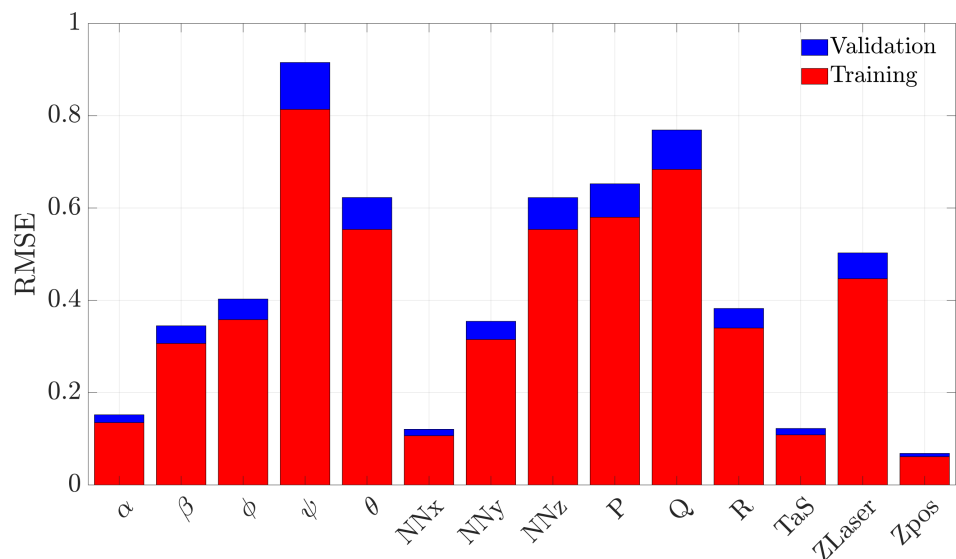
Figure 3. Comparison of the RMSE achieved in training and validation for the 14 linear regression models.

8. FI and FE Performance on the Validation Data (Offline Analysis)

For brevity purposes, only the results of the analysis relative to the air data sensors $\alpha(k)$, $\beta(k)$, and $TaS(k)$ are reported. However, it should be emphasized that the implemented schemes consider the entire set of $n_x = 14$ sensors (residuals); in other words, the FI schemes isolate one among the 14 sensors.

Further, since the main purpose of this research is to compare the performance of the FI and FE techniques, the following analysis was performed assuming an "ideal" failure detection, i.e., the occurrence of a fault is detected as soon as it is injected into a generic time instant $k = k_f$. Clearly, in practice, fault detection is not instantaneous, and a fault detection delay is to be expected before the FI and FE algorithms are activated.

The overall evaluation of the FI and FE of the schemes was performed evaluating the average performance provided by the set of four validation flights (approximately 2.05 h) by injecting the faults in the first sample of each validation flight ($kf = 1$). The analysis was then performed considering, for each sensor, 50 equally-spaced fault amplitudes A_i in (9) in the range $[-A_{M_i}; +A_{M_i}]$. The maximum amplitudes A_{M_i} (see Table 5) were selected empirically so that, for $A_i = A_{M_i}$, the FI algorithms correctly isolate the fault with a percentage greater than 80% (see Section 8.1).

The performance was analyzed using two specific metrics, that is the Fault Isolation Percentage (FIP) and the Relative Fault Reconstruction Error (RFRE). Both metrics are described below.

Table 5. Maximum fault amplitudes.

	$\alpha(k)$ **(deg)**	$\beta(k)$ **(deg)**	$TaS(k)$ **(m/s)**
A_M	2	5	3

8.1. Fault Isolation Percentage

Considering a fault on the i-th sensor of amplitude A_i in the j-th validation flight, the Fault Isolation Percentage (FIP) denoted by $I_{\%_i}$ is defined as the percent ratio over

all the validation flights between the number of samples for which the fault is correctly attributed to the *i*-th sensor and the total number of samples:

$$I_{\%_i}(A_i) = 100 \cdot \sum_{j=1}^{N_{val}} N_{OK_{j,i}}(A_i) \Big/ \sum_{j=1}^{N_{val}} N_{I_j} \tag{48}$$

where:

- N_{val}: number of validation flights.
- $N_{OK_{j,i}}(A_i)$: number of samples for which the fault isolation index $i_F(k)$ correctly isolates the fault on the *i*-th sensor, in validation flight *j*, for fault amplitude A_i.
- N_{I_j}: total number of samples in validation flight *j*.

Figures 4–6 show the FIP for the three sensors as a function of the fault amplitude computed using the $N_{val} = 4$ validation flights for each of the considered techniques.
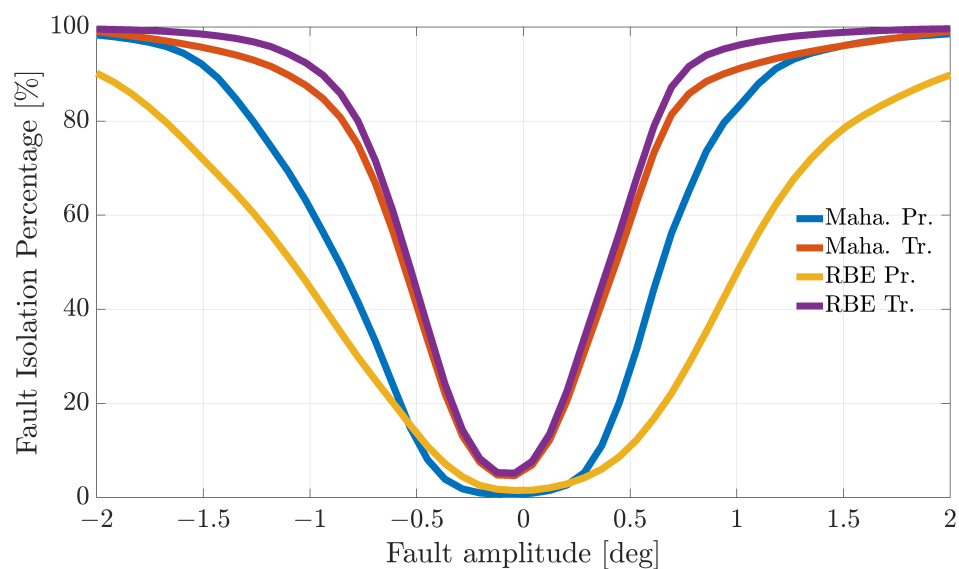
Figure 4. Fault isolation percentage for the $\alpha(k)$ sensor evaluated on the validation flights as a function of the fault amplitude.
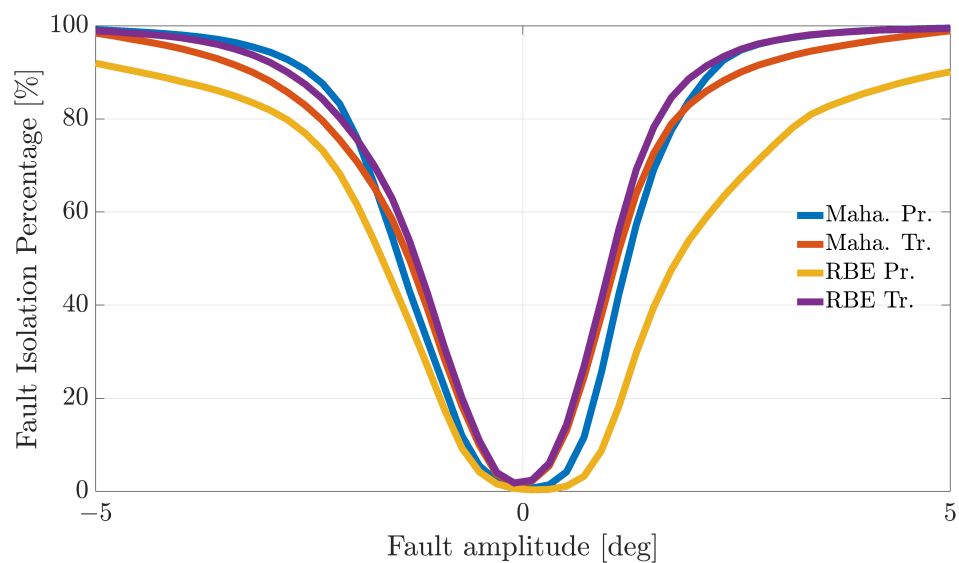
Figure 5. Fault isolation percentage for the $\beta(k)$ sensor evaluated on the validation flights as a function of fault amplitude.
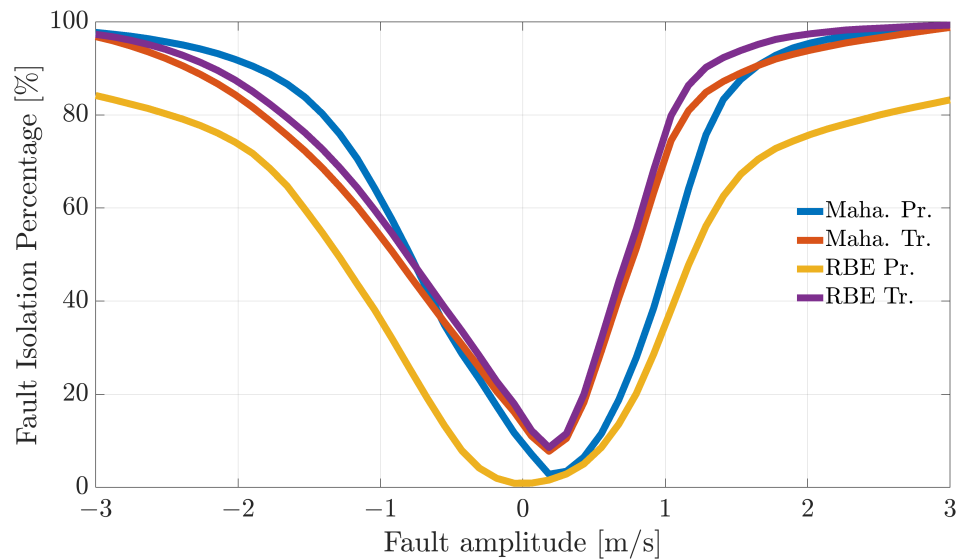
Figure 6. Fault isolation percentage for the *TaS(k)* sensor evaluated on the validation flights as a function of fault amplitude.

The analysis of the plots reveals that all the techniques can guarantee 100% correct fault isolation for large enough fault amplitudes for all the sensors. On the other side, small amplitude faults are often misinterpreted and misclassified. This is not surprising since small-amplitude faults have amplitudes similar to those of modeling errors, making the FI unreliable. The techniques Maha-Trand RBE-Tr based on transformed residuals provide better performance compared to the techniques based on primary residuals. This highlights the fact that the residual transformation based on the diagonalization of the noise covariance matrix is effective at improving FI performance.

8.2. Relative Fault Reconstruction Error

Considering a fault on the *i*-th sensor of amplitude A_i in the *j*-th validation flight, the Relative Fault Reconstruction Error (RFRE) is defined as the percent mean relative amplitude reconstruction error, that is:

$$E_{\%_i}(A_i) = 100 \cdot \sum_{j=1}^{N_{val}} \sum_{k \in S_{E_j}} \left| \frac{\hat{A}_{i_j}(k) - A_i}{A_i} \right| / \sum_{j=1}^{N_{val}} N_{E_j} \tag{49}$$

where:

- A_i: amplitude of the fault on sensor *i*;
- $\hat{A}_{i_j}(k)$: amplitude of the reconstructed fault at sample time *k* for the validation flight *j*.
- S_{E_j}: set of samples in validation flight *j* where the fault is correctly attributed to the *i*-th sensor.
- N_{E_j}: number of samples in the set S_{E_j}.

Figures 7–9 compare the RFRE as a function of the fault amplitude. From the analysis of the figures, it can be observed that all the techniques accurately estimate the amplitude of either positive or negative faults for a large enough fault amplitude; instead, the estimate is not accurate for small amplitude faults. This is not surprising since when the amplitude of the fault has a magnitude of the same order as the estimator modeling error, the relative fault reconstruction is unreliable and inaccurate. For the $E_{\%_i}(A_i)$ index, it is not possible to identify a clear winning approach. Indeed, the RBE-Pr method provided the best results for the $\alpha(k)$ sensor, the worst for the $\beta(k)$ sensor, while for the *TaS(k)* sensor, the performance was comparable.
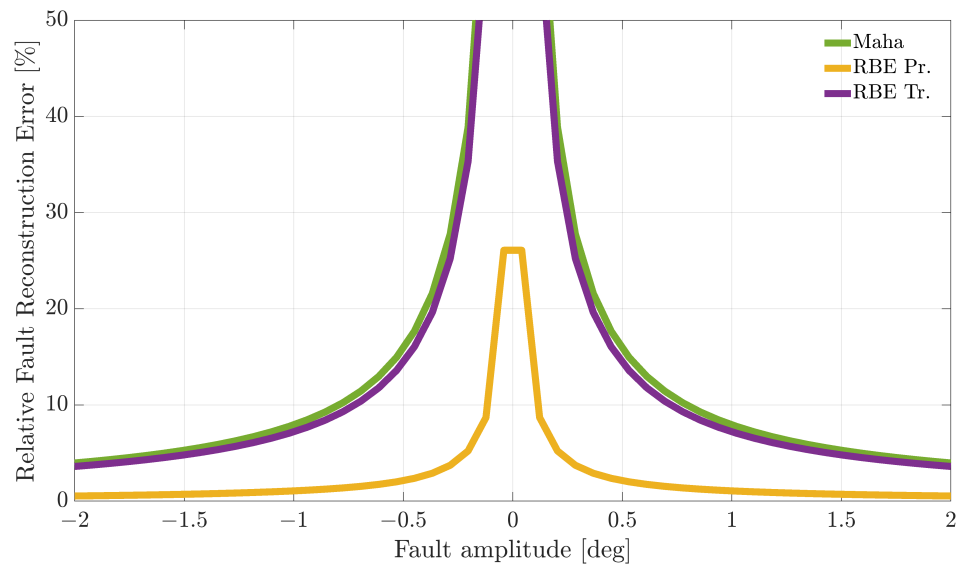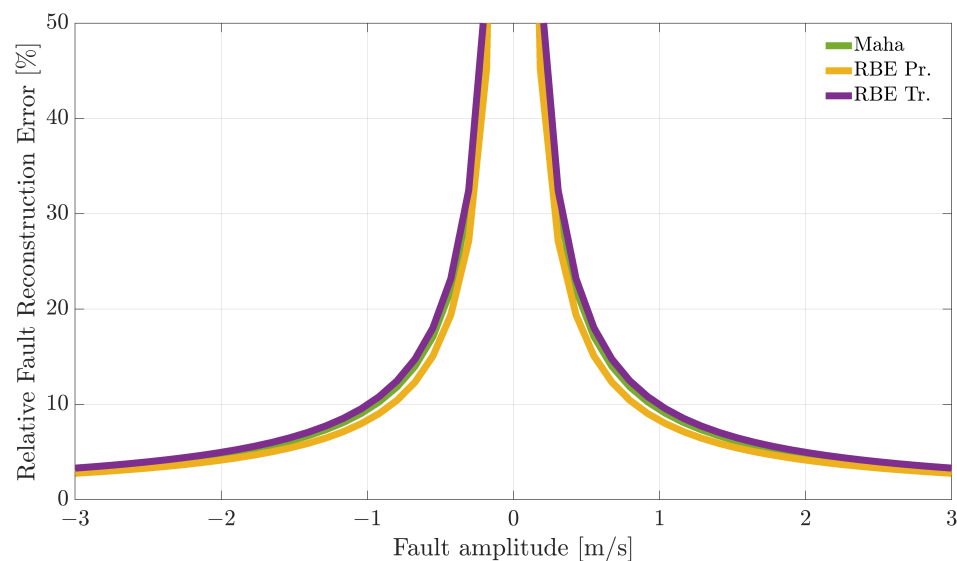
Figure 7. Relative fault reconstruction error for the $\alpha(k)$ sensor evaluated on the validation flights as a function of fault amplitude.
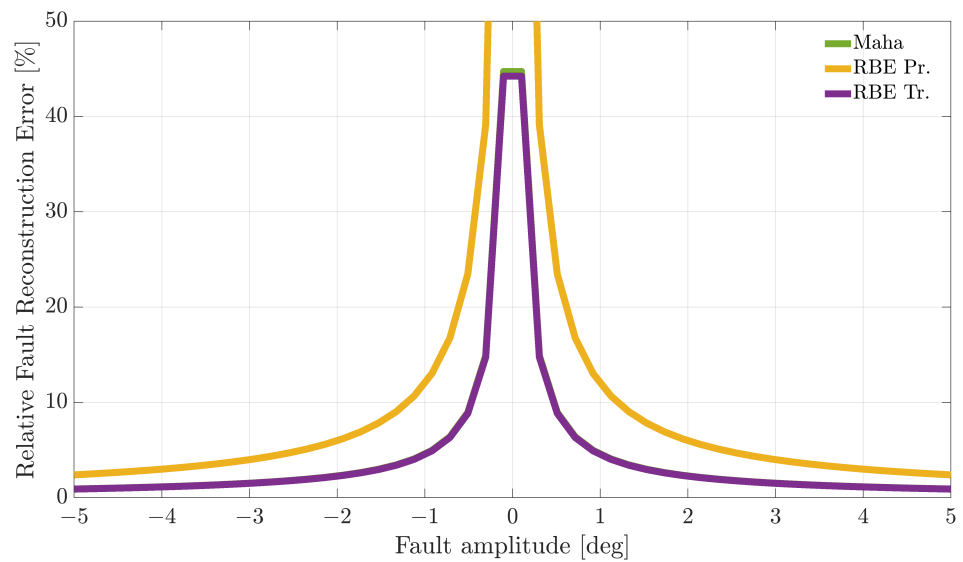

Figure 8. Relative fault reconstruction error for the $\beta(k)$ sensor evaluated on the validation flights as a function of fault amplitude.

Figure 9. Relative fault reconstruction error for the $TaS(k)$ sensor evaluated on the validation flights as a function of fault amplitude.

## 9. FI Performance on Validation Data (Online Analysis)

During the flight, following fault detection, it is critical to quickly isolate the faulty sensors and, next, to take the appropriate reconfiguration actions. In this section, the dynamic response of FI methods based on the recursive BFs introduced in Section 5 is analyzed. To understand the online operation in Figures 10–12, the evolution of the fault beliefs $(p(F_i(k)|e_i(k))$ in (40) for the RBE-Pr technique for a fault injected at t_f equal to 50% of the duration of a validation flight is shown. In the figures, the results for fault amplitudes equal to $\pm A_M$, $\pm A_M/2$, and $\pm$ sample-variance of the residuals in fault fault-free conditions are provided.
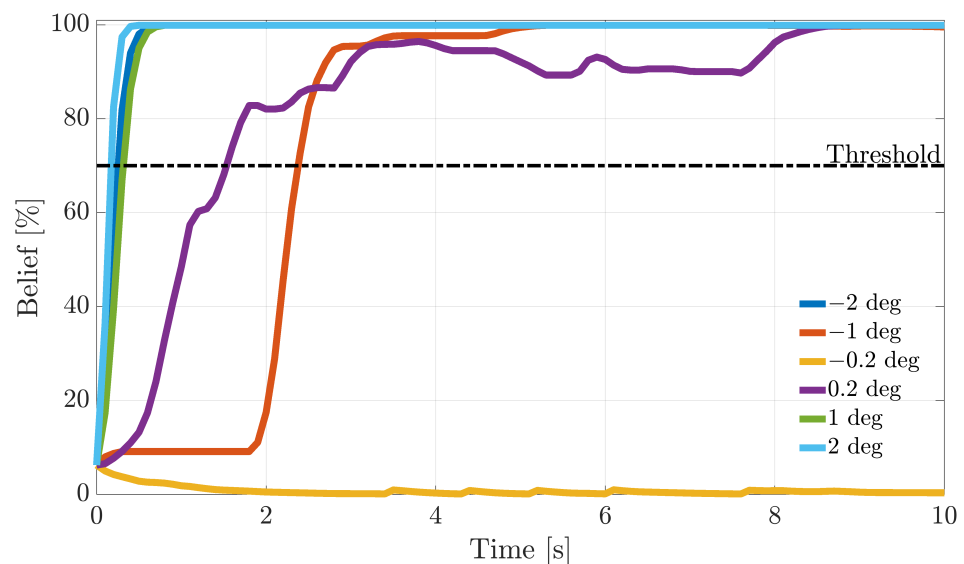
Figure 10. Evolution of the percent fault belief $p[\alpha(k)|e_i(k)]$ following a failure on the $\alpha(k)$ sensor, for different fault amplitudes (RBE-Pr. method).
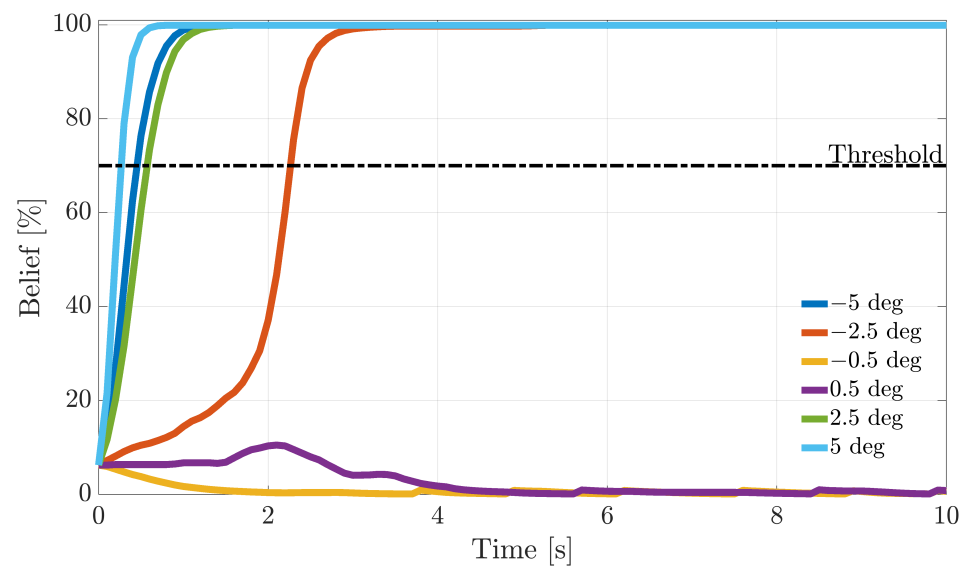
Figure 11. Evolution of the percent fault belief $p[\beta(k)|e_i(k)]$ following a failure on the $\beta(k)$ sensor, for different fault amplitudes (RBE-Pr. method).
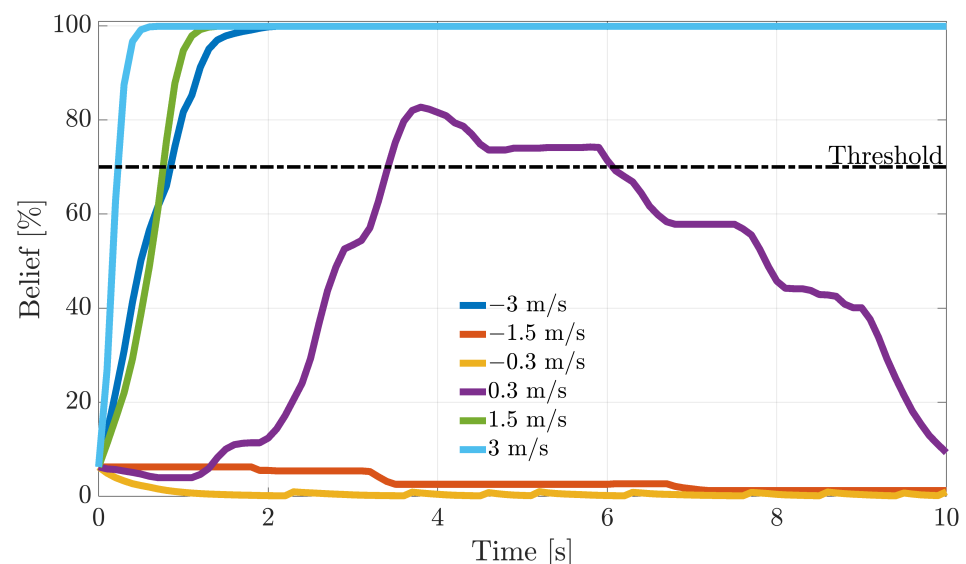
Figure 12. Evolution of the percent fault belief $p[TaS(k)|e_i(k)]$ following a failure on the $TaS(k)$ sensor, for different fault amplitudes (RBE-Pr. method).

It is observed that the fault beliefs grow monotonically for large and intermediate fault amplitudes, reaching in a short time the "fault declaration threshold" (set to 70%) when a fault on the *i*-th sensor is declared. On the other side, for a fault amplitude equal to $\pm$ the sample variance of the modeling error, the fault beliefs in most cases are not strong enough to reach the threshold; thus, the fault cannot be isolated.

Fault Isolation Delay

Fault isolation delay quantifies the (mean) time delay between the generic FD and the instant the FI algorithm reaches, for the first time, a fault belief equal to the predefined "fault declaration threshold". As highlighted in Figures 10–12, this delay depends mainly on the fault amplitude, but may depend also on the in-flight fault instant. To evaluate the online FI performance of the schemes, faults (9) were injected into 10 equally-spaced time

instants k_f for each validation flight. For a fault amplitude A_i on the *i*-th sensor, the (mean) fault isolation delay is defined as:

$$T_{isolation_i}(A_i) = \sum_{j=1}^{N_{val}} \sum_{k_f \in S_{T_j}} T_{isolation_{k_f}}(A_i) / \sum_{j=1}^{N_{val}} N_{T_j} \tag{50}$$

where:

- $T_{isolation_{k_f}}(A_i)$: for a validation flight *j* and for a fault amplitude A_i, denotes the time from the injection of the fault at $k = k_f$ and the time for which the belief associated with the *i*-th sensor reaches for the first time the threshold.
- S_{T_j}: set of equally-spaced instants ($k = k_f$) in the validation flight *j* when the fault (9) is injected in the *i*-th sensor.
- N_{T_j}: number of samples in the set S_{T_j}.

Figures 13–15 show the index (50) as a function of the fault amplitude. For small-amplitude faults, in case the fault is not isolated within an observation window of 10 s, the corresponding fault isolation time is conventionally set to a very large value. Analyzing the figures, it can be observed that Bayesian filtering for all the methods guarantees a reliable declaration of medium-large faults within 3–4 s. For all three sensors, the best performance (smaller $T_{isolation_i}$) was provided by the RBE-Tr. method. Performance degraded quickly for all the methods in the case of small amplitude faults since, in some cases, the fault belief does not reach the fault declaration thresholds within the allowed observation window. Furthermore, in this case, the experimental results confirm that techniques based on transformed directional residuals are more effective (lower mean fault isolation delay) than techniques based on primary directional residuals.
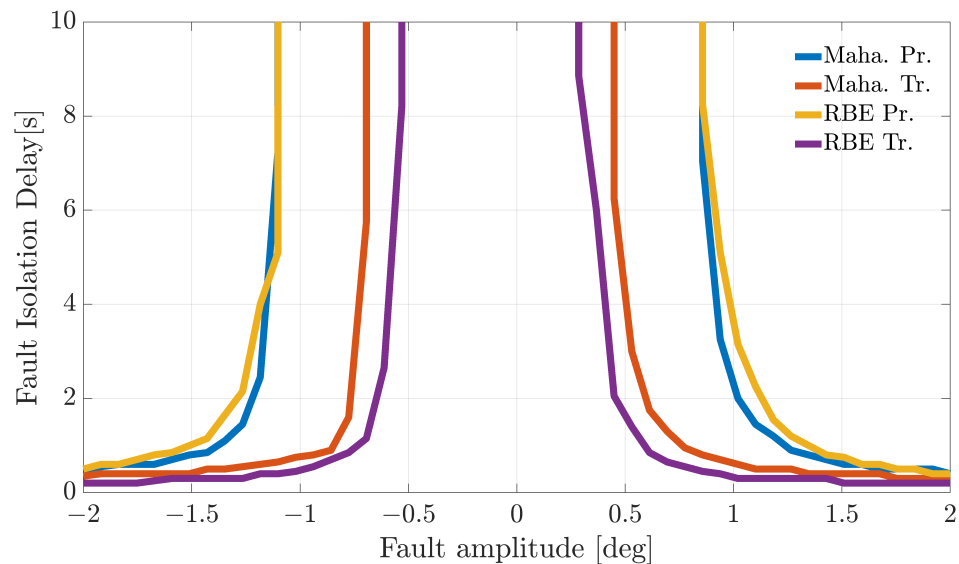
Figure 13. Fault isolation delay for faults on the $\alpha(k)$ evaluated on the validation flights as a function of fault amplitude.
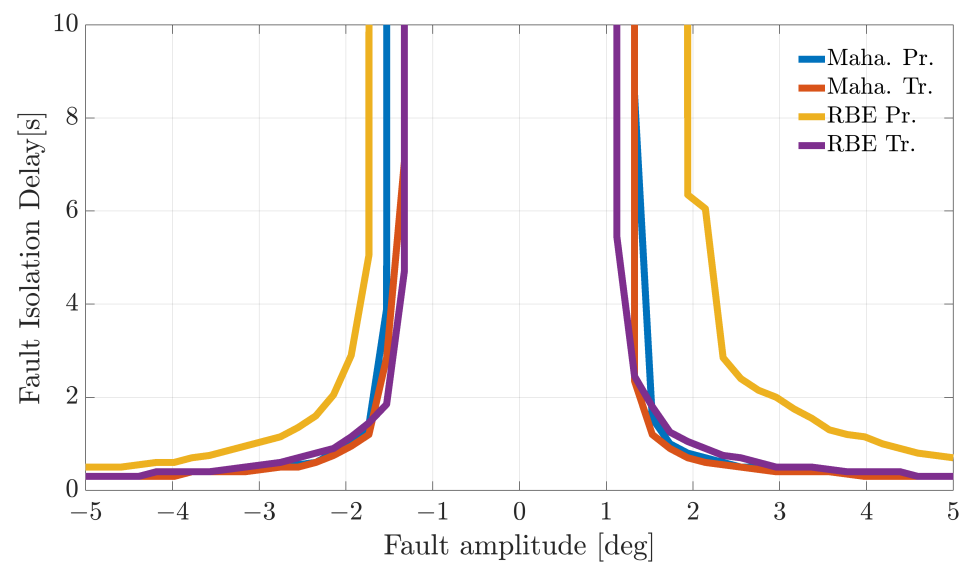
Figure 14. Fault isolation delay for faults on the $\beta(k)$ evaluated on the validation flights as a function of fault amplitude.
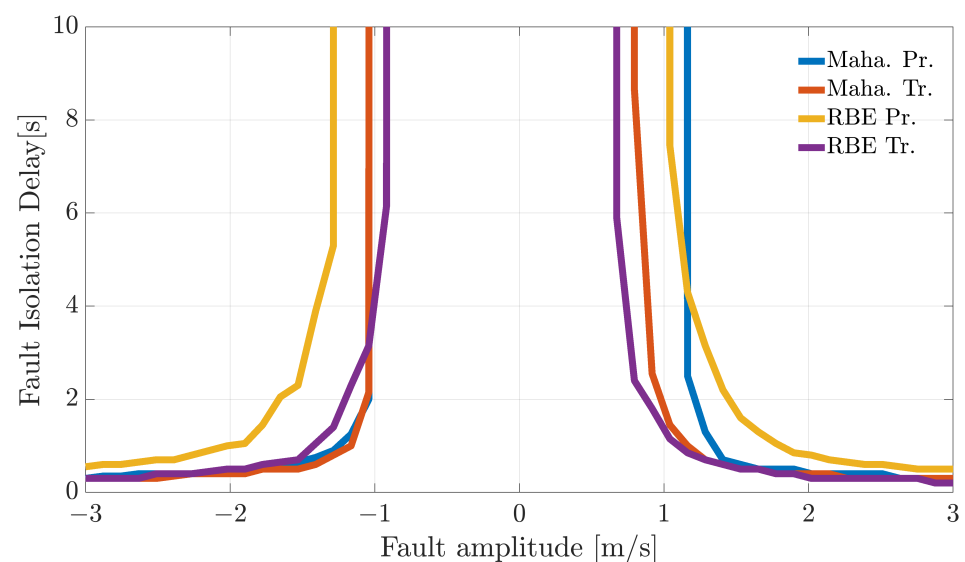
Figure 15. Fault isolation delay for faults on the $TaS(k)$ evaluated on the validation flights as a function of fault amplitude.

10. Conclusions

The purpose of this effort was to compare well-known analytical redundancy-based data-driven techniques for the Fault Isolation (FI) and Fault Estimation (FE) of a set of 14 sensors of a semi-autonomous aircraft. While all these techniques have been shown in the literature to provide close to perfect results using simulated data, only the use of actual experimental data provides the necessary insights and understanding leading to the selection of the best approaches. Specifically, multiple sets of flight data were used to identify linear multivariate models providing a set of primary residuals. Then, residual transformation techniques were applied to generate directional residuals that are robust with respect to the modeling errors. Next, Mahalanobis distance and reconstruction-based methods were used for the FI and the FE. Detailed tests (performed on a set of four validation flights) on the air data sensors showed that the reconstruction-based techniques featuring transformed residuals provide better performance compared to the primary residual-based techniques in terms of the overall fault isolation percentage and fault reconstruction accuracy indices.

In-flight FI was also investigated by applying a bank of recursive Bayesian filters to manage the directional error information online from the 14 sensors. A detailed analysis was conducted by injecting variable amplitude faults at different points throughout the flights. This allowed the estimation of the mean in-flight fault isolation delay. Even for this case, the reconstruction-based methods relying on transformed residuals provided the best performance.

All the considered FI and FE methods are data-driven and were designed based on actual flight data. Therefore, the schemes do not require a priori knowledge of detailed aircraft modeling for their implementation and can be easily returned regularly with updated flight data. Although this is undoubtedly a useful aspect, it is also worth noting, as for any data-based technique, that the reliability of the results depends heavily on the completeness of the available data, which must be representative of all operating conditions. In order to address this issue, all our models were derived by taking multi-flight data rich in maneuvers (as opposed to cruise steady-state conditions), both in the design and validation phases. The quantitative results provide a clear picture of the requested design effort as far as the achievable performance using well-known FI and FE schemes.

Author Contributions: Conceptualization, N.C., M.R.N. and M.L.F.; methodology, M.R.N. and M.L.F.; software, N.C.; validation, N.C.; formal analysis, N.C.; investigation, N.C. and G.C.; resources, M.L.F.; data curation, M.R.N. and M.L.F.; writing—original draft preparation, N.C.; writing—review and editing, M.R.N. and M.L.F.; visualization, N.C. and G.C.; supervision, M.R.N. and M.L.F.; project administration, M.L.F.; funding acquisition, G.C. and M.L.F. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare that there are no conflicts of interest in performing this research.

Abbreviations

The following abbreviations are used in this manuscript:

Maha. Pr.	Mahalanobis distance-based technique with Primary residuals
Maha. Tr.	Mahalanobis distance-based technique with Transformed residuals
RBE Pr.	Reconstruction-based technique with Primary residuals
RBE Tr.	Reconstruction-based technique with Transformed residuals

References

1. Ding, S. *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*; Springer: Berlin/Heidelberg, Germany, 2008. [CrossRef]
2. Gertler, J. Analytical Redundancy Methods in Fault Detection and Isolation-Survey and Synthesis. *IFAC Proc. Vol.* **1991**, *24*, 9–21. [CrossRef]
3. Simani, S.; Fantuzzi, C.; Patton, R. *Model-Based Fault Diagnosis in Dynamic Systems Using Identification Techniques*; Springer: Berlin/Heidelberg, Germany, 2003. [CrossRef]
4. Isermann, R. Process fault detection based on modeling and estimation methods—A survey. *Automatica* **1984**, *20*, 387–404. [CrossRef]
5. Basseville, M. Detecting changes in signals and systems—A survey. *Automatica* **1988**, *24*, 309–326. [CrossRef]
6. Gertler, J. *Fault Detection and Diagnosis in Engineering Systems*; Routledge: London, UK, 2017.
7. Patton, R.J.; Frank, P.M.; Clark, R.N. *Issues of Fault Diagnosis for Dynamic Systems*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
8. Gao, Z.; Cecati, C.; Ding, S.X. A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis with Model-Based and Signal-Based Approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [CrossRef]
9. Blanke, M.; Kinnaert, M.; Lunze, J.; Staroswiecki, M. *Diagnosis and Fault Tolerant Control*; Springer: Berlin/Heidelberg, Germany, 2016.

10. Ding, S. Optimal fault detection and estimation: A unified scheme and least squares solutions. *IFAC-PapersOnLine* **2018**, *51*, 465–472. [CrossRef]

11. Li, L.; Ding, S.; Peng, X. Optimal Observer-based Fault Detection and Estimation Approaches for T-S Fuzzy Systems. *IEEE Trans. Fuzzy Syst.* **2020**. [CrossRef]

12. Martinez-Guerra, R.; Mata-Machuca, J.L. *Fault Detection and Diagnosis in Nonlinear Systems*; Springer: Berlin/Heidelberg, Germany, 2016.

13. Sobhani-Tehrani, E.; Khorasani, K. *Fault Diagnosis of Nonlinear Systems Using a Hybrid Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009; Volume 383.

14. Zhang, Q.; Zhang, X.; Polycarpou, M.M.; Parisini, T. Distributed sensor fault detection and isolation for multimachine power systems. *Int. J. Robust Nonlinear Control* **2014**, *24*, 1403–1430. [CrossRef]

15. Isermann, R. Model-based fault-detection and diagnosis—Status and applications. *Annu. Rev. Control* **2005**, *29*, 71–85. [CrossRef]

16. Ding, S. *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]

17. Yin, S.; Ding, S.X.; Haghani, A.; Hao, H.; Zhang, P. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *J. Process Control* **2012**, *22*, 1567–1581. [CrossRef]

18. Mrugalski, M. *Advanced Neural Network-Based Computational Schemes for Robust Fault Diagnosis*; Springer: Berlin/Heidelberg, Germany, 2014.

19. Bocaniala, C.D.; Palade, V. Computational intelligence methodologies in fault diagnosis: Review and state of the art. In *Computational Intelligence in Fault Diagnosis*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1–36.

20. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [CrossRef]

21. López-Estrada, F.R.; Méndez López, L.; Santos-Ruiz, I.; Valencia-Palomo, G. Detección de fallas en vehículos aéreos no tripulados mediante señales de orientación y técnicas de aprendizaje de máquina. *Revista Iberoamericana de Automática e Informática Industrial RIAI* **2021**. [CrossRef]

22. Schaefer, R. *Unmanned Aerial Vehicle Reliability Study*; Office of the Secretary of Defense: Washington, DC, USA, 2003.

23. Goupil, P. AIRBUS state of the art and practices on FDI and FTC in flight control system. *Control Eng. Pract.* **2011**, *19*, 524–539. [CrossRef]

24. Johnson, D.M. A review of fault management techniques used in safety-critical avionic systems. *Prog. Aerosp. Sci.* **1996**, *32*, 415–431. [CrossRef]

25. Marzat, J.; Piet-Lahanier, H.; Damongeot, F.; Walter, E. Model-based fault diagnosis for aerospace systems: A survey. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **2012**, *226*, 1329–1360. [CrossRef]

26. Farsoni, S.; Simani, S. *Robust Fault Diagnosis and Fault Tolerant Control of Wind Turbines: Data-Driven and Model-Based Approaches*; Scholars' Press: Riga, Latvia, 2016.

27. Chu, E.; Gorinevsky, D.; Boyd, S. *Detecting Aircraft Performance Anomalies from Cruise Flight Data*; AIAA Infotech@Aerospace: Atlanta, Georgia, 2010. [CrossRef]

28. Li, L.; Gariel, M.; Hansman, R.J.; Palacios, R. Anomaly detection in onboard-recorded flight data using cluster analysis. In Proceedings of the 2011 IEEE/AIAA 30th Digital Avionics Systems Conference, Seattle, WA, USA, 16–20 October 2011; pp. 4A4-1–4A4-11. [CrossRef]

29. Dani, M.C.; Freixo, C.; Jollois, F.; Nadif, M. Unsupervised anomaly detection for Aircraft Condition Monitoring System. In Proceedings of the 2015 IEEE Aerospace Conference, Big Sky, MT, USA, 7–14 March 2015; pp. 1–7. [CrossRef]

30. Li, L.; Das, S.; John Hansman, R.; Palacios, R.; Srivastava, A.N. Analysis of Flight Data Using Clustering Techniques for Detecting Abnormal Operations. *J. Aerosp. Inf. Syst.* **2015**, *12*, 587–598. [CrossRef]

31. Lu, P.; Van Eykeren, L.; van Kampen, E.J.; de Visser, C.; Chu, Q. Double-model adaptive fault detection and diagnosis applied to real flight data. *Control Eng. Pract.* **2015**, *36*, 39–57. [CrossRef]

32. Fravolini, M.L.; Napolitano, M.R.; Core, G.D.; Papa, U. Experimental interval models for the robust Fault Detection of Aircraft Air Data Sensors. *Control Eng. Pract.* **2018**, *78*, 196–212. [CrossRef]

33. Fravolini, M.L.; Del Core, G.; Papa, U.; Valigi, P.; Napolitano, M.R. Data-Driven Schemes for Robust Fault Detection of Air Data System Sensors. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 234–248. [CrossRef]

34. Tecnam P92 Webpage. 2020. Available online: https://www.tecnam.com/aircraft/p92-echo-mkii/ (accessed on 22 December 2020).

35. Gertler, J.J.; Kunwer, M.M. Optimal residual decoupling for robust fault diagnosis. *Int. J. Control* **1995**, *61*, 395–421. [CrossRef]

36. Basseville, M. Information criteria for residual generation and fault detection and isolation. *Automatica* **1997**, *33*, 783–803. [CrossRef]

37. Hu, Y.; Gertler, J. Design of Directional Residuals for Optimal Testability. *IFAC Proc. Vol.* **2002**, *35*, 131–136. [CrossRef]

38. Alcala, C.F.; Qin, S.J. Reconstruction-based contribution for process monitoring. *Automatica* **2009**, *45*, 1593–1600. [CrossRef]

39. Frisk, E.; Nielsen, L. Robust residual generation for diagnosis including a reference model for residual behavior. *Automatica* **2006**, *42*, 437–445. [CrossRef]

40. Varrier, S.; Koenig, D.; Martinez, J.J. Robust fault detection for Uncertain Unknown Inputs LPV system. *Control Eng. Pract.* **2014**, *22*, 125–134. [CrossRef]

41. Witczak, M.; Buciakowski, M.; Puig, V.; Rotondo, D.; Nejjari, F. An LMI approach to robust fault estimation for a class of nonlinear systems. *Int. J. Robust Nonlinear Control* **2016**, *26*, 1530–1548. [CrossRef]

42. Balzano, F.; Fravolini, M.; Napolitano, M.; d'Urso, S.; Crispoltoni, M.; Core, G. Air Data Sensor Fault Detection with an Augmented Floating Limiter. *Int. J. Aerosp. Eng.* **2018**, *2018*, 1072056. [CrossRef]

43. Leondes, C.T. *Techniques in Discrete and Continuous Robust Systems*; Academic Press: Cambridge, MA, USA, 1996.

44. Mahalanobis, P.C. On the generalized distance in statistics. *Proc. Natl. Inst. Sci.* **1936**, *2*, 49–55.

45. De Maesschalck, R.; Jouan-Rimbaud, D.; Massart, D. The Mahalanobis distance. *Chemom. Intell. Lab. Syst.* **2000**, *50*, 1–18. [CrossRef]

46. Qin, S.J. Survey on data-driven industrial process monitoring and diagnosis. *Annu. Rev. Control* **2012**, *36*, 220–234. [CrossRef]

47. Cartocci, N.; Costante, G.; Napolitano, M.R.; Valigi, P.; Crocetti, F.; Fravolini, M.L. PCA Methods and Evidence Based Filtering for Robust Aircraft Sensor Fault Diagnosis. In Proceedings of the 2020 28th Mediterranean Conference on Control and Automation (MED), Saint-Raphaël, France, 16–19 June 2020; pp. 550–555. [CrossRef]

48. Yang, F.; Xiao, D. Progress in Root Cause and Fault Propagation Analysis of Large-Scale Industrial Processes. *J. Control Sci. Eng.* **2012**, *2012*, 478373. [CrossRef]

49. Massoumnia, M. A geometric approach to the synthesis of failure detection filters. *IEEE Trans. Autom. Control* **1986**, *31*, 839–846. [CrossRef]

50. Hu, Y.; Gertler, J. Design of optimal directional residuals for linear dynamic systems. *IFAC Proc. Vol.* **2003**, *36*, 245–250. [CrossRef]

51. Lou, X.C.; Willsky, A.S.; Verghese, G.C. Optimally robust redundancy relations for failure detection in uncertain systems. *Automatica* **1986**, *22*, 333–344. [CrossRef]

52. Särkkä, S. *Bayesian Filtering and Smoothing*; Cambridge University Press: Cambridge, MA, USA, 2013; Volume 3.

53. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.

54. Ge, Z.; Zhang, M.; Song, Z. Nonlinear process monitoring based on linear subspace and Bayesian inference. *J. Process Control* **2010**, *20*, 676–688. [CrossRef]

55. Zheng, Y.; Mao, S.; Liu, S.; Wong, D.S.; Wang, Y. Normalized Relative RBC-Based Minimum Risk Bayesian Decision Approach for Fault Diagnosis of Industrial Process. *IEEE Trans. Ind. Electron.* **2016**, *63*, 7723–7732. [CrossRef]

56. Zhu, J.; Ge, Z.; Song, Z. Distributed Parallel PCA for Modeling and Monitoring of Large-Scale Plant-Wide Processes with Big Data. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1877–1885. [CrossRef]

57. Zhou, W.; Yang, W.; Wang, Y.; Zhang, H. Generalized Reconstruction-Based Contribution for Multiple Faults Diagnosis with Bayesian Decision. In Proceedings of the 2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS), Enshi, China, 25–27 May 2018; pp. 813–818. [CrossRef]

58. Basseville, M.; Nikiforov, I.V. *Detection of Abrupt Changes—Theory and Application*; Prentice Hall, Inc.: Upper Saddle River, NJ, USA, 1993; p. 550.

59. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A Survey. *ACM Comput. Surv.* **2009**, *41*. [CrossRef]

60. Draper, N.; Smith, H. *Applied Regression Analysis*, 3rd ed.; A Wiley-Interscience Publication; Wiley: New York, NY, USA, 1998.

61. The Mathworks Inc. *MATLAB—MathWorks*; The Mathworks Inc.: Natick, MA, USA, 2020.