On the Use of the Indirect Test Strategy for Lifetime Performance Monitoring of RF Circuits

Hassan El Badawi, Florence Azaïs, Serge Bernard, Mariane Comte, Vincent Kerzérho, François Lefèvre

On the use of the indirect test strategy for lifetime performance monitoring of RF circuits

H. El Badawi [1,2], F. Azaïs [1], S. Bernard [1], M. Comte [1], V. Kerzerho [1], F. Lefevre [2]
[1] LIRMM, Univ. Montpellier, CNRS, 161 rue Ada, Montpellier, France
[2] NXP Semiconductors, 2 Espl. Anton Phillips, 14000 Caen, France

***Abstract*— On-line performance monitoring of RF integrated circuits throughout their operating lifetime is a promising way to enhance their reliability. This paper explores the feasibility of adapting the indirect test strategy to implement on-line RF performance monitoring. After stating the principle of the proposed solution, we consider the fundamental requirements necessary to adapt the indirect test strategy. Finally, a proof of concept is established through a practical case of study by monitoring the power level delivered by the RF transmitter of a wireless microcontroller: hardware measurement results demonstrate the potential of this approach.**

Keywords*— *RF integrated circuits, reliability, on-line monitoring, indirect test, machine-learning

1 Introduction

The market of RF devices has experienced an accelerated growth in the last decades, particularly with the deployment of various telemetry and remote monitoring applications, inducing a competitive high production context where quality is a key factor. While production test guarantees the quality of the devices once manufactured and sent to the client, dedicated methods for on-line performance monitoring are required to ensure their reliability once deployed, especially for devices used in critical-safety applications such as fire detection, entrance access control, smart metering… By allowing early detection of device performance degradation, these monitoring methods could warn the user in time to avoid a complete system failure.

Unlike digital devices where reliability issues have been extensively studied and a large variety of on-line testing techniques have been developed [1], very few works concern analog/RF performance monitoring. In [2], the authors study the design of an adaptive checker for concurrent error detection based on common mode signal analysis. A real time estimation is investigated in [3] in order to accurately monitor a performance by capturing the variation of the distortion performance. In [4] the use of an embedded temperature sensor is proposed, in [5] a current-based circuit for in-field monitoring is presented. In this paper, we take a different approach, which aims to tailor the indirect test strategy to the monitoring objective.

The indirect test strategy has initially been proposed in the context of analog/RF circuit production test to reduce testing costs [6,7]. The basic principle of this strategy is briefly described hereafter to facilitate understanding of its adaptation as a means of on-line performance monitoring. Assuming that process variations affect both the device RF parameters, which are conventionally measured inducing high testing costs, and Indirect Measurements (IMs), which can be obtained at a much lower cost, the objective is to establish the underlying correlation between these two spaces in order to predict the RF performances from the sole indirect measurements. Since the relationship between these two sets of parameters is complex and cannot be expressed by a simple analytical function, machine-learning algorithms are used, involving two phases in the classical implementation of the indirect testing strategy, as illustrated in Figure 1. The first is the initial learning phase, in which a set of training devices is used to build regression models that relate indirect measurements to RF performance; for this set of training devices, both conventional RF measurements and low-cost IM measurements are performed and feed the machine-learning algorithms. The obtained models are further operated in the second phase, namely the production test phase, for any new device to be tested; from there, only the low-cost measurements are performed, while the RF performances are predicted through the regression models. The final binning as good or faulty device is done on the basis of the predicted values of the RF performances.

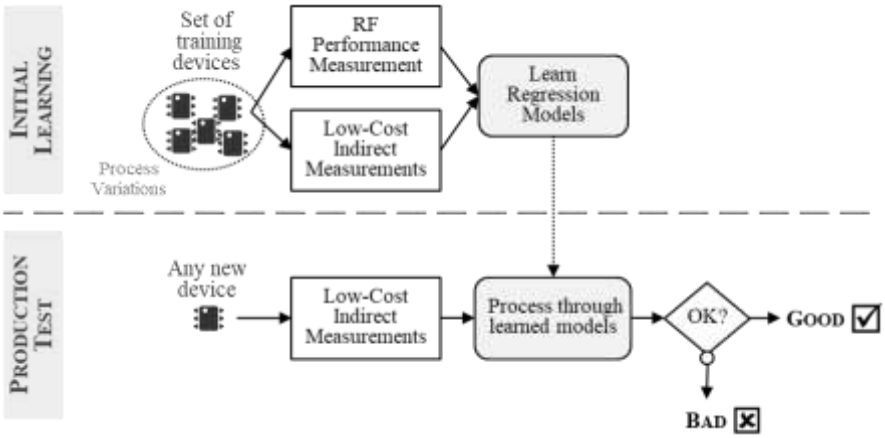

Fig. 1. Indirect test synopsis

Many publications investigate the efficiency of the indirect test strategy depending on different factors, such as the choice of the type of regression model, the selection of the most pertinent indirect measurements, the use of embedded sensors, or the implementation of an adaptive test flow among others. In [8], a complete overview of these studies is presented. The use of the indirect test strategy to implement Built-In Self-Test (BIST) has also been investigated in [9-12]. In [13,14], the authors explore a post-manufacturing calibration of analog circuits. Our objective in this paper is to exploit the indirect test strategy in order to develop an on-line performance monitoring scheme for RF circuits. A preliminary work has been presented in [15]. This paper is an extended version of that work that includes a comprehensive analysis of the collected dataset alongside with a detailed description of the embedded measurement process, especially with respect to the use of averaging to overcome the limitations of the embedded ADC.

The paper is organized as follows. The principle of the proposed strategy is introduced in Section II and the main requirements needed to implement the approach are discussed. The test vehicle used as a case study is presented in Section III and the dataset collected in the measurement campaign is then analyzed in Section IV. Section V is dedicated to the initial learning phase and details the elaboration of the prediction model. Section VI is dedicated to the practical implementation of the on-line monitoring process and hardware experimental results obtained on the case study are presented. Finally, Section VII concludes the paper.

2 Adaptation of the Indirect Test Strategy for On-Line Performance Monitoring

2.1 Principle

The principle of the proposed strategy for on-line performance monitoring is illustrated in Figure 2. As in the classical indirect test implementation, it involves a preliminary learning phase in which the mapping between a given circuit performance and some indirect measurements is established through the construction of a regression model. The main difference is that the learning set should include not only devices affected by process variations but also devices representative of the main wear-out failure mechanisms susceptible to occur during the circuit life. It is therefore recommended that the learning set includes devices that have been subjected to accelerated life tests or burn-in. Note that such tests are normally part of the qualification process performed before launching mass-volume production. The only requirement here is that not only the conventional performance measurements have to be performed, but also the indirect measurements.

Once the learning phase is over, mass production can start. Every new manufactured device undergoes a production test; devices that do not comply with the specifications are rejected (this test can be implemented using either a conventional approach or an indirect test solution). Before shipping good

devices to the customers, there is then an additional step which consists in storing within the IC the original value of the performance that will be monitored and the values of the model coefficients determined in the previous phase.

Finally, once the device is deployed in its application, the on-line monitoring process can be triggered at any time. It involves the embedded measurement of the selected IMs and the embedded computation of the performance value. In this process, the device predicts its own performance variation based on the model established during the learning phase and the values stored within the IC at production time; a flag alert is raised if the predicted performance variation exceeds a predefined threshold.

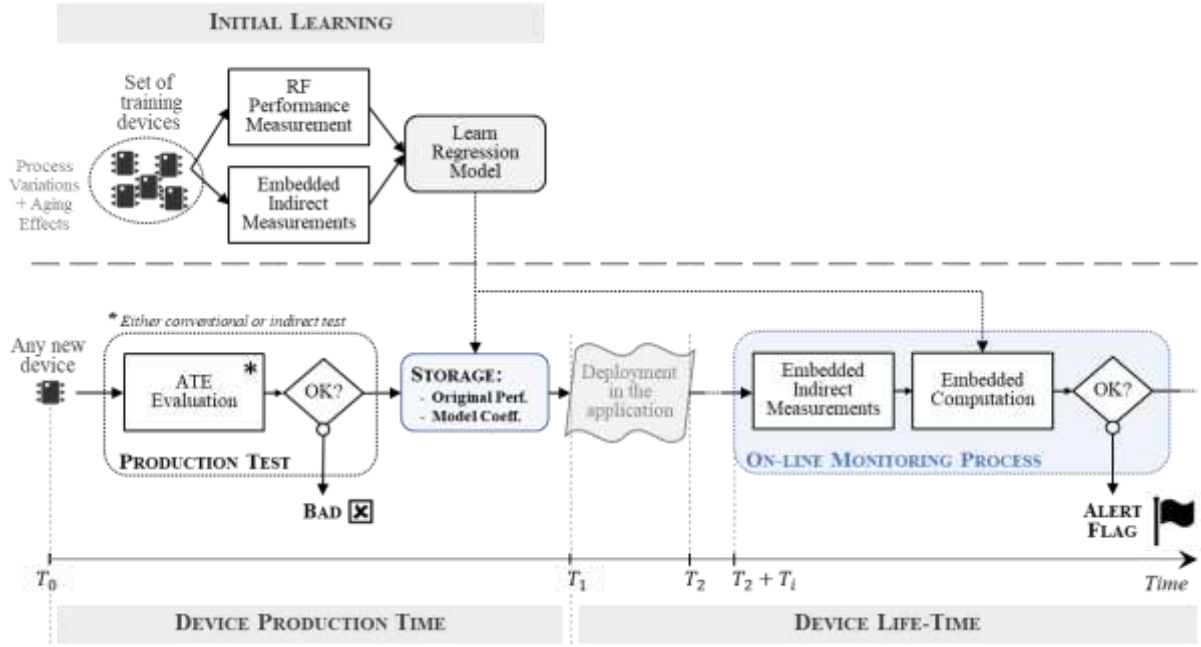

Fig. 2. Principle of on-line RF performance monitoring based on the indirect test strategy

2.2 Hardware Requirements

To perform the embedded prediction, a number of hardware resources are obviously necessary: (i) a dedicated infrastructure in order to access internal nodes or structures involved in the indirect measurements, (ii) digitization resources to convert the measured analog values into the digital domain, (iii) a non-volatile memory to store and fetch the coefficients of the established regression model and finally, (iv) a processing unit to perform the computation of performance prediction. The main requirements on these resources are discussed hereafter.

Note that all these resources do not necessarily have to be embedded within the circuit itself; some of them might be available within the system in the application, for instance the digitization resources and/or the processing unit. However, the choice of the pertinent indirect measurements has to be done by the IC provider at the design stage and the minimum requirement is that the circuit is equipped with a dedicated infrastructure allowing the end user to access these parameters. In case digitization and processing are not included within the circuit, an application note that details how to exploit the indirect measurements should also be provided.

2.2.1 Indirect Measurements

Assembling a comprehensive set of pertinent indirect measurements is a keystone to achieve efficient implementation of an indirect test strategy for production testing. The same is of course true for an on-line performance monitoring solution based on the indirect test strategy. However, a specific requirement for an on-line monitoring process is that all the considered IMs must have the possibility to

be measured on-chip, and preferably with a simple measurement infrastructure. Hence, not all types of indirect measurements considered as potential candidates in the context of production testing are applicable in the context of an on-line monitoring solution. For example, the standard DC tests usually applied on external pins are achieved with the help of the ATE resources and cannot be easily performed by the device itself. In the same way, the possibility of changing the test conditions, e.g. the power supply voltage, is not an issue during production testing but it is much trickier to implement when the device is deployed in the field. In this context, the most natural candidates are DC voltages on internal nodes and DC signatures delivered by built-in sensors accessible through an internal analog test bus.

Furthermore, the main motivation behind the on-line performance monitoring strategy is to observe and detect any performance deterioration induced by aging effects. The most important aging phenomena observed today in nanometric technologies are hot carrier injection (HCI), time dependent dielectric breakdown (TDDB), bias temperature instability (BTI) and electromigration (EM) [16]. These aging phenomena affect not only the digital parts but also the analog/RF parts [17]. They result in internal variations and alterations of the integrated circuit characteristics, such as a shift of the threshold voltage, which can heavily impact the value of the bias voltage of an amplifier, or a shorter gate-oxide breakdown lifetime among other effects [18]. Hence, in order to perform lifetime performance monitoring, it is highly desirable that the set of IM candidates include indirect measurements sensitive to these wear-out mechanisms. Due to practical constraints, this aspect is not considered for the case study used in this paper.

2.2.2 Digitization

Once the indirect measurements and the test infrastructure allowing to access these measurements are defined, the following step is the choice of the digitization resources. Indeed, the measured analog values must be converted into digital values in order to be further processed for the computation of the embedded prediction. It is essential that the quantization error introduced by this conversion does not significantly affect the accuracy of the computed prediction. The choice and the design of the digitization resources is therefore an important aspect.

The choice should take into consideration the characteristics of the different indirect measurements used for performance prediction. Indeed, it is likely that the indirect measurements cover a large voltage range while the variation range of each indirect parameter might be small. The digitization resources have to cope with this diversity without comprising the conversion accuracy. In particular, the voltage resolution of the ADC used to perform the conversion of a given indirect parameter must be much smaller than the variation range of this parameter. The voltage resolution of an ADC is equal to its measurement range (or full-scale range), divided by the number of quantization levels, i.e. 2^n, where n is the number of bits of the converter. To ensure an appropriate voltage resolution, it is therefore possible to play either on the measurement range or on the number of bits.

In this context, several options can be considered for the design of the required digitization resources, i.e. (i) a single high-resolution ADC with a large measurement range that covers the complete variation range of all indirect measurements, (ii) a single medium-resolution ADC with a programmable measurement range that can be adapted to groups of indirect measurements with a similar order of magnitude in the variation range, or (iii) several low-resolution ADCs, each one with a fixed measurement range perfectly adapted to the variation range of one indirect measurement. The retained solution obviously strongly depends on the case study and will be a tradeoff between the required silicon area and the conversion accuracy.

2.2.3 Memory and Arithmetic Unit

A regression model is defined by (i) the function that relates the indirect measurements to the predicted performance and, (ii) a set of coefficient values that parametrizes the regression function. In order to implement an embedded prediction, it is therefore necessary to have memory as well as arithmetic resources. The memory resources are used to store the value of the coefficients established during the learning phase. These values obviously need to be permanently stored in the circuit or the system, which implies the use of a non-volatile memory. Alongside the memory, an arithmetic unit must

be included in the circuit or system to perform the calculations defined by the established regression model.

It is important to highlight that performance monitoring of a device in the field is an auxiliary option to improve the reliability of the system, but is not the main core of the application. Therefore, the additional circuitry required to implement the embedded prediction must be minimized. Moreover, the processing time must be minimized in order to maintain the normal operation of the system without disruption. Hence, it is essential to reduce the number of required operations; the choice of the regression model type plays an important role in this aspect.

In the classical implementation of the indirect test strategy for production test, the model accuracy is crucial since binning of devices as good or bad circuits is realized based only on the results of the prediction. To achieve high accuracy, models usually implemented are non-linear models such as Multi-Adaptive Regression Spline (MARS) or Support Vector Machine (SVM) models, or even more sophisticated models based on ensemble methods [19]. However, the use of these types of model is problematic in the context of on-line performance prediction. Indeed, they involve the storage of a substantial number of coefficients as well as the computation of specific non-linear functions that cannot be easily implemented with a standard arithmetic unit. Embedded performance computation based on such models would therefore be consuming both in terms of memory resources and processing time, which is a strong drawback.

An alternative way is to lean on less accurate but easier to implement models, such as Multiple Linear Regression (MLR) models. Indeed, such models are simple models that compute the predicted value ($\hat{y}$) as a weighted sum of the input parameters (x_i):

$$\hat{y} = c_0 + \sum_{i=1}^{m} c_i \cdot x_i. \qquad (1)$$

where c_i are the model coefficients.

These models involve only basic arithmetic operations and a limited number of coefficients ($m + 1$ for a model with m parameters). Of course, the simplicity of implementation comes at the cost of lower model accuracy. However, in the context of on-line performance monitoring, the accuracy constraint is not as strong as in the context of production test since the objective is just to monitor whether the performance has experienced a degradation and to quantify the order of magnitude of this degradation. This is therefore the choice that is used in this study.

Finally, note that an interesting option has been suggested in [20] to enhance the performance of linear prediction models by creating an enriched set of feature candidates from the initial available feature set. More precisely, they have proposed a Python library to generate non-linear features ($\log(x)$, $\exp(x)$, $|x|$, $1/x$, x^2, x^3) and combine pairs of features with various operators ($+, -, *$); feature selection is then applied on this enriched set to build linear prediction models. Such an approach is promising in the context of embedded performance prediction, since some non-linear transformations or interactions have the potential to be implemented with few memory resources and only basic arithmetic operations, implying low processing time. We will therefore consider this option in this study.

3 Case Study

3.1 Test Vehicle: RF Transceiver (NXP JN518x)

The test vehicle considered as a case study is an ultra-low power wireless microcontroller developed by NXP Semiconductors supporting Zigbee 3.0 and Thread networking stacks to facilitate the development of home automation, smart lighting and wireless sensor network applications. It includes a 2.4GHz IEEE 802.15.4 compliant transceiver and a comprehensive mix of analog and digital peripherals, as well as an Arm Cortex-M4 core and embedded Flash and RAM memory. It also includes an embedded test infrastructure that involves an analog DC bus allowing to probe various internal nodes of the transceiver (11 nodes) and a 12-bit ADC. The internal nodes might be connected either to two General Purpose Input Output (GPIO) pins, or to the embedded ADC. The general architecture of this product is summarized in Figure 3.

Our objective is to detect a possible degradation of the power level delivered by the transmitter, which is an essential performance of the system. This product has been selected for this study because it is equipped with all the necessary hardware to implement embedded performance prediction. Specifically, the indirect measurements (voltage on internal nodes of the transceiver) will be performed using the test infrastructure and digitized by the embedded ADC. They will be transferred to the CPU where they will be processed using the model equation determined during the initial learning phase along with the model coefficients and the original performance value stored in the Flash memory during the production. The predicted variation will be compared to a predefined threshold and a flag alert will be issued on a GPIO pin if the value is below the threshold.

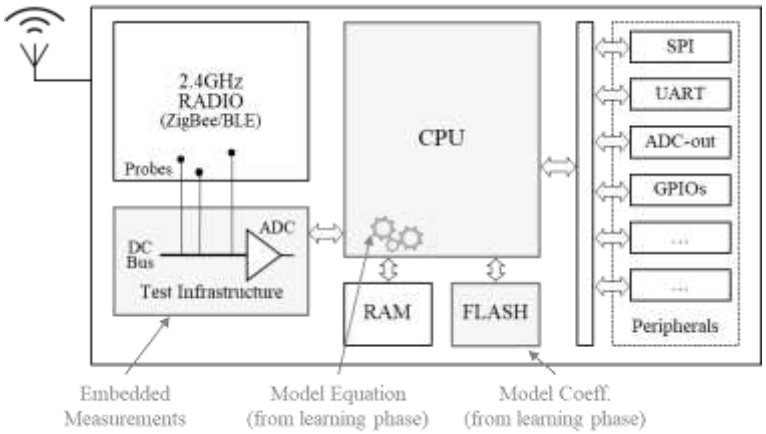

Fig. 3. High-level block diagram of the test vehicle

At this point, it is important to highlight that the implementation of an on-line monitoring process was not considered during the design phase of this product. Hence, no specific structures were embedded in the product with respect to the prediction of the transmitted power level, nor with respect to the impact of aging effects. In the same way, the ADC available in the circuit has not been specifically designed for conversion of the embedded measurements. Our objective with this case study is therefore to study the feasibility of implementing on-line performance monitoring based on the indirect test strategy, and we do not expect a high accuracy from the prediction model. The case study should be considered as a proof-of-concept rather than a validation of the efficiency that can be achieved with this approach.

3.2 Dataset Collection: Measurement Campaign

A measurement campaign has been performed to collect data in order to build a regression model that maps indirect measurements to the level of transmitted power. Ideally, data should be collected on circuits that present a degradation of the power level due to aging effect. However, because it is a new product and the mass-production is not yet launched, we do not have this opportunity. In this context, the solution retained to collect data representative of a power level degradation was to act on two internal configuration registers of the transmitter. Each one of these registers has four configuration bits (16 configurations). By playing on both registers, we therefore have 256 possibilities to operate the transmission block of the device under a different configuration. One of these configurations corresponds to the nominal configuration (C172), and the others are used to emulate a variation of the transmitted power level.

Practically, test data were collected from four ICs on Advantest V93K ATE. Each IC has been operated under the 256 possible configurations and the transmitted power level has been measured using the ATE RF resources, for each configuration. The DC voltage on the 11 internal nodes that can be accessed through the analog test bus have also been measured by the ATE (DC30 to DC40 indirect measurements). In total, the dataset contains 12,288 measurements resulting from 1,024 observations, as depicted in Figure 4.

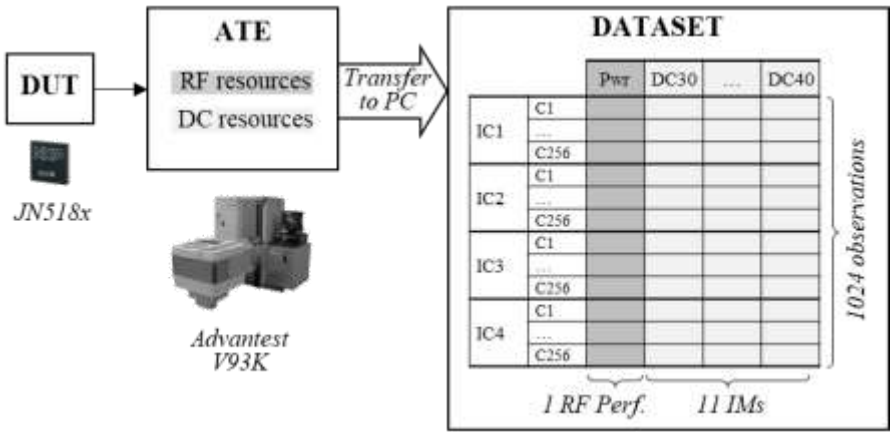

Fig. 4. Measurement campaign

4 Dataset Analysis

The main characteristics of the collected dataset are summarized in Table I, which reports the nominal value observed under the nominal configuration (mean value over the 4 ICs) as well as the absolute and relative variation range observed over the others 255 configurations, for the transmitted power level and the 11 indirect measurements. Several comments arise from the analysis of this table.

TABLE I. CHARACTERISTICS OF COLLECTED DATA (1024 OBSERVATIONS OVER 4 ICS)

	Nominal Value *(mean over the 4 ICs)*	**Variation Range**	**Relative Variation Range**
Power Level	11.42 dBm	[-3.34 dB; +1.82 dB]	[-29%; +16%]
DC30	1.065 V	[-17.9 mV; +15.2 mV]	[-1.7%; +1.4%]
DC31	1.062 V	[-23.8 mV; +19.3 mV]	[-2.2%; +1.8%]
DC32	1.058 V	[-19.4 mV; +17.2 mV]	[-1.8%; +1.6%]
DC33	0.133 V	[-5.0 mV; +4.3 mV]	[-3.8%; +3.2%]
DC34	1.126 V	[-15.6 mV; +17.0 mV]	[-1.4%; +1.5%]
DC35	1.9 mV	[-0.4 mV; +0.6 mV]	[-20%; +30%]
DC36	1.129 V	[-15.5 mV; +17.2 mV]	[-1.4%; +1.5%]
DC37	1.4 mV	[-0.2 mV; +0.3 mV]	[-16.7%; +21%]
DC38	1.128 V	[-15.6 mV; +17.2 mV]	[-1.4%; +1.5%]
DC39	0.592 V	[-6.6 mV; +6.4 mV]	[-1.1%; +1.1%]
DC40	1.074 V	[-8.4 mV; +4.0 mV]	[-0.8%; +0.3%]

First regarding the transmitted power level, it can be observed that playing on the value of the internal configuration registers indeed permits to obtain a variation of this performance, but in a limited range. Indeed, the maximum variation from the nominal value is only of -3.34 dB, which corresponds to a degradation of -29%. We have verified that this variation is mainly related to the setting of the internal configuration registers rather than the variability from one IC to another. This is confirmed by the fact that the difference between two ICs in the value of the transmitted power level remains below 0.43 dB over all the configurations, which corresponds to a variability below 3.7%. So, even if this is not the most favorable condition to build a regression model, we expect that the observed variation is significant enough.

Regarding the indirect measurements, two main comments can be drawn. First, it should be highlighted that the nominal value can be very different from one IM to another: seven IMs have a high nominal value between 1V and 1.13V (DC30, DC31, DC32, DC34, DC36, DC38, DC40), one IM has a medium nominal value around 0.6V (DC39), one IM has a low nominal value around 0.1V (DC33), and two IMs have a very low nominal value below 2mV (DC35, DC37). However, in all cases, the

observed variation is small, i.e. only few tens of mV for most of the IMs and even only few hundreds of µV for two IMs (DC35 and DC37). This might cause accuracy issue for the embedded measurement. The second comment is that, despite the small variation range for all IMs, the situation differs if we look at the relative variation. More specifically, most of the indirect measurements present a low relative variation, i.e. below 4%, which is much lower than the one observed on the transmitted power level. However, two IMs (DC35 and DC37) exhibit a significant relative variation that is the same range than the one observed for the transmitted power level.

Here again, it is useful to establish whether this variation is related to the setting of the internal configuration registers or to the variability from one IC to another. For this, we have looked on the one hand at the mean of the relative variation over the four ICs in each configuration, and on the other hand at the mean of the relative variation over the 256 configurations for each IC. Results are depicted in Figures 5 and 6 respectively, for the eleven indirect measurements (each dot in figure 5 corresponds to the mean value computed over the four IC in a given configuration, whereas each dot in figure 6 corresponds to the mean value computed over the 256 configurations for a given IC.

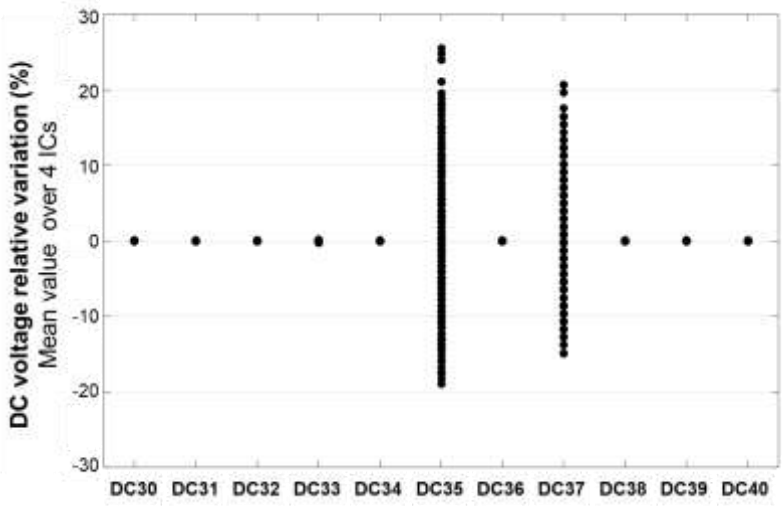

Fig. 5. Mean of the relative variation over the four ICs for the eleven IMs

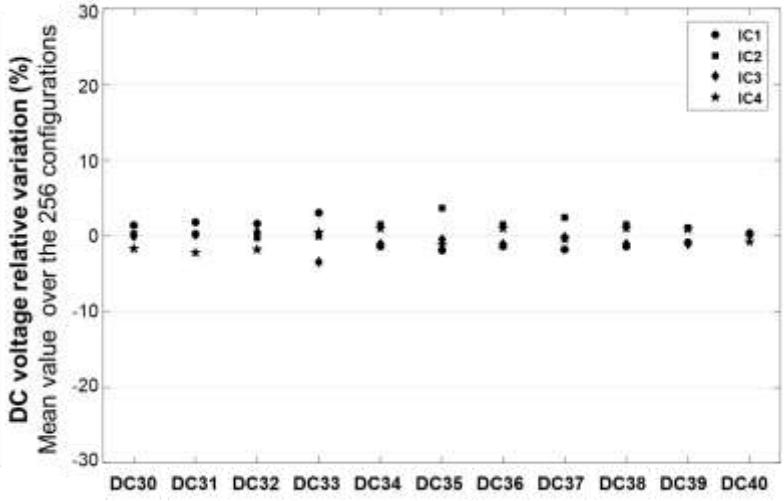

Fig. 6. Mean of the relative variation over the 256 configurations for the eleven IMs

Figure 5 clearly shows that only DC35 and DC37 are actually impacted by the circuit configuration with a relative variation that exceeds 20%, all the other indirect measurements presenting an almost constant value with a relative variation lower than 0.03% (dots corresponding to the 256 configurations

are almost superimposed on the 0 value on the y-axis for these IMs). In contrast, Figure 6 shows that there is no significant difference between the indirect measurements with respect to circuit variability, since they all present a comparable variation range, which remains contained below 5%.

This analysis reveals that only two out of the eleven available indirect measurements exhibit a variation related to the internal register configuration (which has an influence on the transmitted power level) higher than the variation from one IC to another induced by the manufacturing process. It is clear that this situation is not the best context to implement an efficient solution for an on-line performance monitoring based on an indirect test strategy. Here again, it is important to keep in mind that the available IMs have not been specifically defined for the context of this study. Hence, they are not necessarily relevant and it is likely that more pertinent IMs could have been defined during the design phase. Despite the weakness of this dataset, the expectation is that we can build a regression model that predicts the power level with a reasonable accuracy and establish a proof-of-concept of the proposed strategy.

5 Initial Learning: Model Elaboration

In this section, the choice of the model type is commented and the procedure for IM selection and model construction is detailed. Results on the accuracy of the retained model to predict the power level variation on IC4 using measurements performed on the ATE are also presented.

5.1 Choice of Model Type

As mentioned in section II, the choice of a simple MLR model allows to easily implement the required calculations and to minimize the number of coefficients that have to be stored within the circuit. A classical approach to build such a model is based on the use of Sequential Forward Selection (SFS) procedure in order to select pertinent IMs [21]. The procedure starts by building an MLR model for each available IM and selecting the IM that generates the model with the minimum prediction error (lowest *RMSE* score). At the second iteration, an MLR model is built for each pair of IMs that includes the previously selected IM; the pair that gives the best model is then selected. The process then continues with triplets and so on.

This procedure has been implemented in this work. However, in order to improve the accuracy of the MLR model, IM selection has been performed on an enriched space of candidates that includes not only the original IMs but also non-linear transformations of these IMs as well as interactions between pairs of IMs.

Regarding the non-linear transformations, some transformations such as $1/x$ and x^2 can be implemented at low-cost because they require only a limited number of elementary arithmetic operations. In contrast, other transformations such as $\log(x)$, $\sqrt{x}$ and $\exp(x)$ would require much longer processing times. Indeed, their exact computation is not feasible with elementary arithmetic operations; instead, numerical algorithms that involve many elementary arithmetic operations have to be used to compute an approximation. In this study, we have considered only transformations that can be implemented at low-cost, i.e. $1/x$ and x^2. Regarding the interaction between IMs, all combinations of pairs of IMs using the four elementary operators $(+, -, *, /)$ can be easily implemented. Interactions using $(+)$ and $(-)$ operators are intrinsically present in the model; therefore, only interactions using $(*)$ and $(/)$ operators have been considered in this study.

Globally, with the considered non-linear transformations and interactions, an enriched space of 209 candidates has been generated from the original space of 11 IM candidates.

5.2 Model Construction

The full dataset of 1024 observations has been partitioned into training and validation sets: the training set includes data collected on IC1, IC2 and IC3 (768 observations) while the validation set is composed of data collected on IC4 (256 observations). The SFS procedure has been applied on the training set and models including between 1 and 5 features have been built to predict the transmitted power level $\hat{P}$.

Results are summarized in Figure 7, which reports the evolution of the model accuracy in terms of Root-Mean-Square Error (RMSE) with respect to the number of features used by the model, for models constructed both on the original and enriched spaces. This figure shows that that there is no significant improvement in the model accuracy by using more than three features. This figure also shows the benefit brought by the enrichment of the candidate space, with a reduction of the RMSE score of about 0.025dB, which corresponds to an accuracy improvement of 6%. Note that although there is a one-to-one correspondence between the number of features and the number of IMs in case of a model constructed on the original space, this is not true in case of a model constructed on the enriched space since a selected feature might involves an interaction between two IMs.

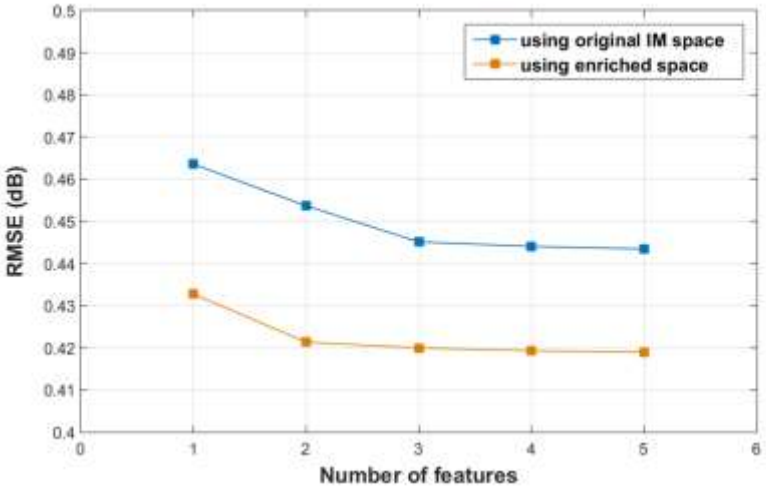

Fig. 7. Model accuracy vs. the number of selected features

The retained solution for all following experiments is the use of an MLR model build with three features selected from the enriched candidate space with:

- Feature 1: $DC39/DC35$
- Feature 2: $1/DC37^2$
- Feature 3: $DC30 * DC35$

The model is defined by the following equation:

$$\hat{P} = a_0 + a_1 * \left(\frac{DC39}{DC35}\right) + a_2 * \left(\frac{1}{DC37^2}\right) + a_3 * (DC30 * DC35) \qquad (2)$$

where $\hat{P}$ is the predicted power level and a_0, a_1, a_2, and a_3 are the model coefficients that have to be stored in the Flash memory of the circuit.

This model necessitates the measurement of four IMs, i.e. $DC30$, $DC35$, $DC37$ and $DC39$. Note that $DC35$ and $DC37$ are logically involved in the model since they are the two IMs that present a significant variation under the different configurations, but it is the interactions of $DC35$ with other IMs and a non-linear transformation of $DC37$ that are exploited (a model build using only the raw version of $DC35$ and $DC37$ actually exhibits a poor accuracy).

Finally, since the objective is to detect a possible degradation of the power level delivered by the transmitter, the power level variation $\widehat{\Delta P}_i^j$ is computed with:

$$\widehat{\Delta P}_i^j = \hat{P}_i^j - P_0^j \qquad (3)$$

where $\hat{P}_i^j$ is the power level predicted for device j in configuration i and P_0^j refers to the original performance measured at production time, which corresponds in this experiment to the power level measured on the ATE for device j in the nominal C172 configuration.

Results obtained on both the training and validation sets regarding power level variation are illustrated in Figure 8. In this figure, each point corresponds to one circuit measured in one of the 256

configurations; the value on the x-axis corresponds to the actual variation determined from direct RF power level measurements and the value on the y-axis corresponds to the predicted variation computed using the prediction model and the value of the four indirect measurements $DC30$, $DC35$, $DC37$ and $DC39$. It can be observed that the regression model performs acceptably well considering the different shortcomings, i.e. the use of a simple linear regression model, the fact that the IMs were not specifically defined for power level prediction and the limited size of the dataset. It can also be observed that there is no discrepancy between prediction results on the ICs of the training set and the IC of the validation set, which is an unseen IC for the model.

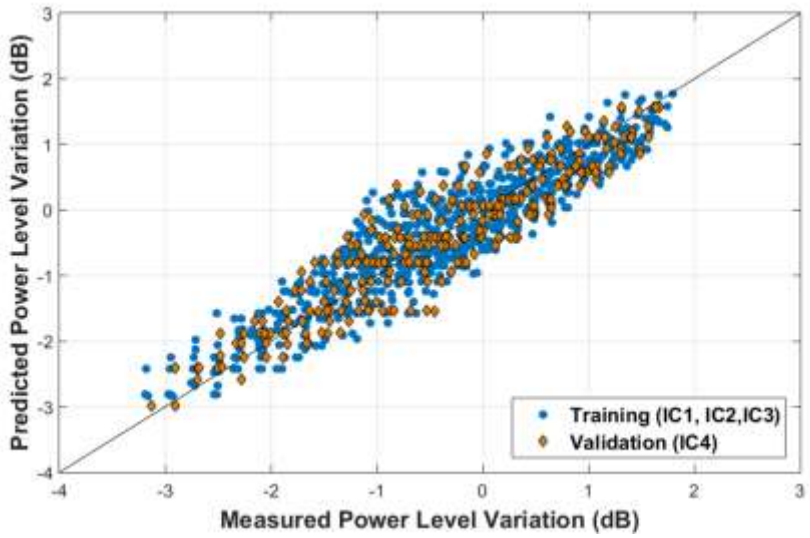

Fig. 8. Predicted power level variation vs. measured power level variation

Regarding the prediction accuracy, results are illustrated in Figure 9, which gives the normalized distribution of the prediction error; numerical results are summarized in Table II. Results show that the prediction error is well-centered on 0 and that it follows a nearly Gaussian distribution. The standard deviation is around 0.42dB, which corresponds to a rms accuracy of about 3.7% and the maximum error is around 1.3dB. We can therefore expect that any power degradation stronger than this will be definitely detected by the on-line monitoring process. For lower degradation, detection might be possible but is not guaranteed since it falls within the model uncertainty.

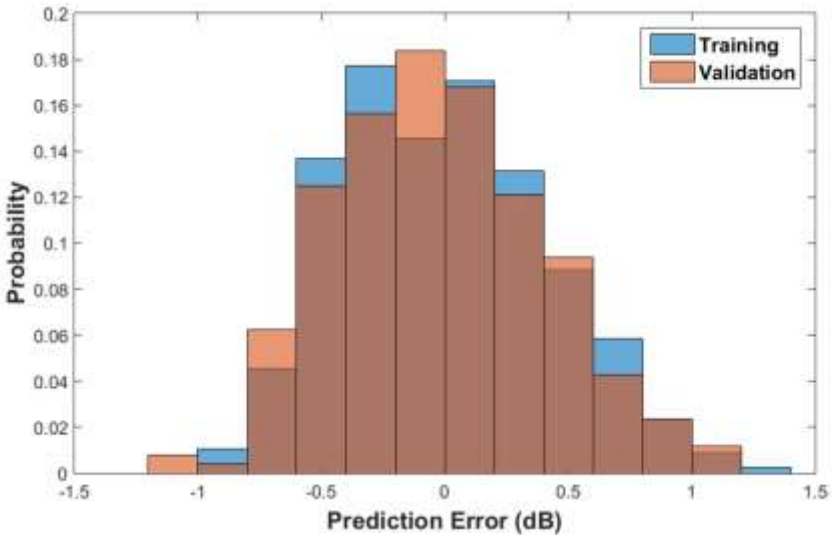

Fig. 9. Normalized distribution of the prediction error

TABLE II. STATISTICS OF THE PREDICTION ERROR OVER THE 256 CONFIGURATIONS

	All configurations		
	Mean	**Std. Dev.**	**Max**
Prediction error -Training	0.00 dB	0.42 dB	1.30 dB
Prediction error - Validation	-0.02 dB	0.42 dB	1.19 dB

6 Embedded Prediction

This section is dedicated to the practical implementation of the on-line monitoring process. More precisely, the regression model elaborated in the previous section from ATE measurements on IC1, IC2, and IC3 is exploited to perform embedded prediction of the power level variation on IC4. The experimental setup is first described and hardware measurement results are then presented.

6.1 Experimental Setup and Product Programming

To emulate the operation of the circuit within its application environment, the circuit is mounted on a development board provided by NXP Semiconductors. MCUXpresso Integrated Development Environment (IDE) is used to develop the software code on a PC. A stand-alone debug probe is then used to download the code within the product and debug the firmware through a JTAG interface. This experimental setup is illustrated in Figure 10.

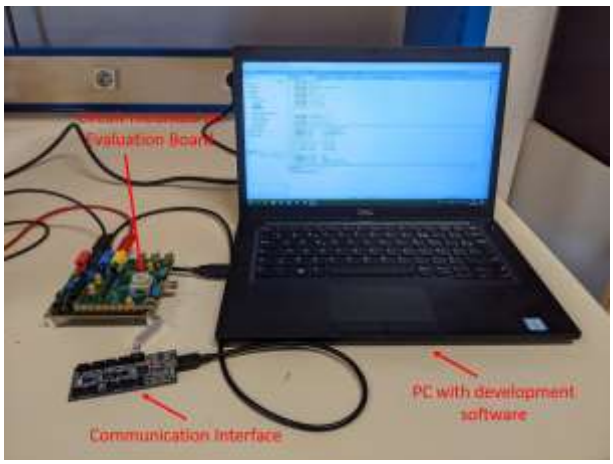

Fig. 10. Illustration of the experimental setup

A specific code dedicated to the on-line performance monitoring process has been developed. The flowchart of this code is illustrated in Figure 11. The code is divided in three main stages:

- In the first stage, the product is initialized and the Test Mode is started. In this mode, the transmitter is activated and delivers a modulated RF signal at 2.4GHz corresponding to internally-generated DSSS sequences.
- Once the circuit is ready, the second stage dedicated to the embedded measurements is launched. In this stage, we loop through the selected indirect measurements in order to measure and store their value. More precisely, for each selected indirect measurement, the DC bus is configurated, acquisition is realized by the ADC and the digitized value is stored in the SRAM memory.
- Finally, the last stage is dedicated to the embedded computation and the verification of the performance. In this stage, the predicted power level degradation is computed based on the model equation established during the learning, the measured IM values stored in the SRAM memory, the model coefficients and the original power level stored in the Flash memory. This value is then simply compared to a predefined threshold and a flag is raised if the value is below the threshold, indicating that the product has experienced a degradation.

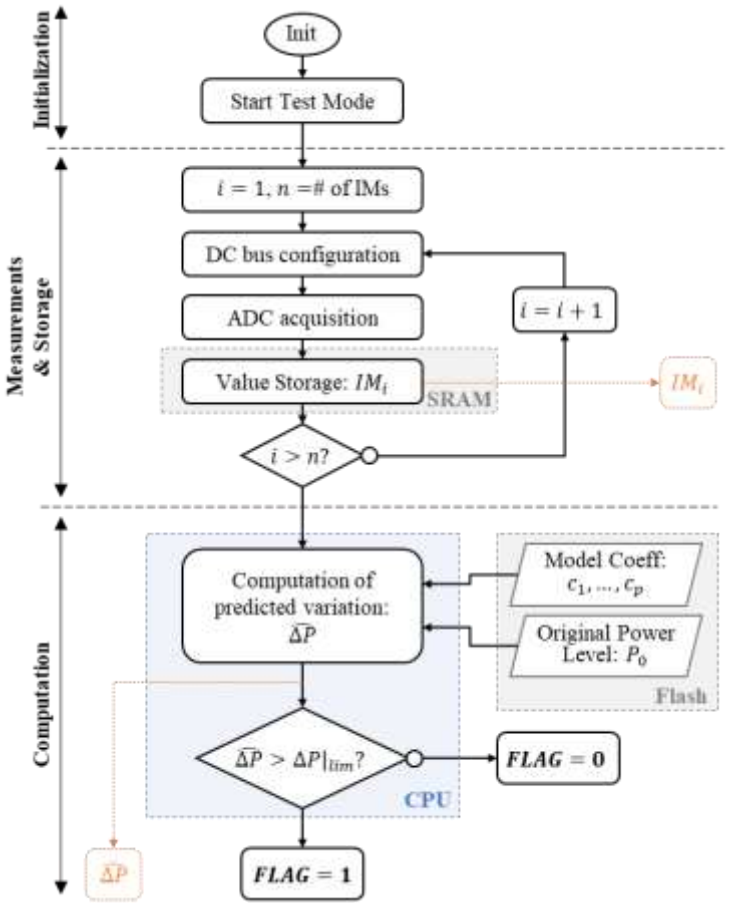

Fig. 11. Flowchart of the on-line monitoring process

Note that for validation purpose, a number of outputs that correspond to different stages have been included during the execution of the code, i.e. the digitized value of the embedded measurements, the predicted power level and the status of the flag alert. These outputs can be transferred to the PC via the USB interface and extracted with the help of the IDE. In the final version of the code, the only output will be the alert flag.

6.2 Initial Results

The developed code has been launched for two configurations of the product, i.e. the nominal configuration (C172) and a degraded one (C241). For each configuration, 100 runs have been performed in order to analyze the repeatability of the embedded prediction results. Results are summarized in Figure 12, which illustrates the power level variation predicted in each configuration, for the 100 runs.

This figure reveals that the embedded measurement and computation process leads to a large dispersion of the predicted power level variation over the 100 runs. More precisely, predicted values are spread in the range [-2.31dB; +1.96dB] for the nominal configuration and [-8.69dB; 0.61dB] for the degraded one, whereas predicted values are expected to be contained in the range [-1.26dB; +1.26dB] for the nominal configuration and [-4.39dB; -1.87dB] for the degraded one (from $\pm 3\sigma$ model uncertainty, symbolized by the dotted lines located around the ideal line in Figure 12). Obviously, such a large dispersion that exceeds the model uncertainty does not permit to ensure the detection of a power level degradation. We presume that this large dispersion comes from measurement repeatability/accuracy problems.

Several elements support this assumption. First, DC35 and DC37, which are essential indirect measurements in the model, have a very low DC value in the range of few millivolts (cf. Section 4). Measurement of such small values is sensitive to noise and therefore might be subject to a measurement

repeatability issue. Moreover, the variation induced by the circuit configuration is also very small, i.e. the maximal deviation from the nominal value is only of 0.34mV. Correct evaluation of such a small deviation necessitates high measurement precision. Practically, the ADC available for the case study can be programmed only for two different measurement ranges, i.e. from 0 to 3.6V and from 0 to 0.9V; this second measurement range is chosen for the measurement of DC35 and DC37. The ADC has a resolution of 12 bits, which means that one LSB corresponds to 0.22mV. A deviation of 0.34mV therefore corresponds to a difference of 1.5 LSB, which means that only 2 bits out of the 12 bits of the ADC are actually exploited. Obviously, such a low equivalent resolution has a strong impact on the measurement accuracy and therefore on the prediction error.

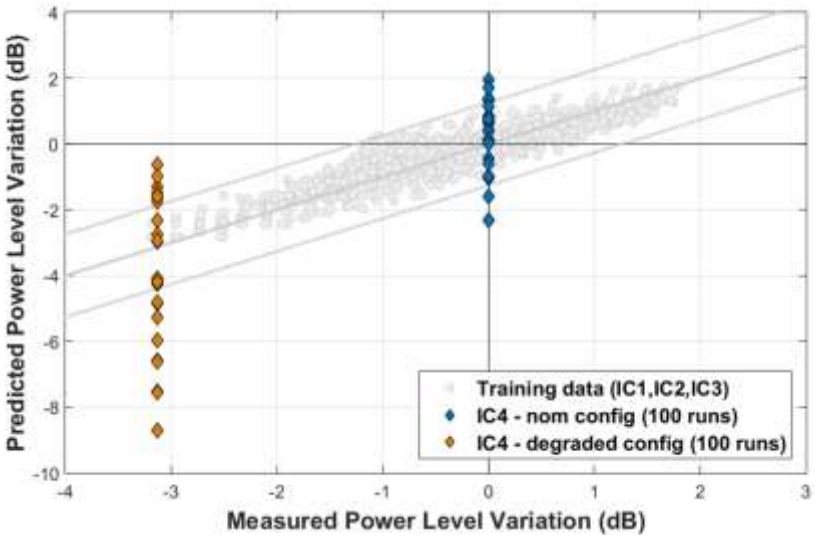

Fig. 12. Embedded prediction of power level variation for IC4 in two configurations

6.3 Use of averaging

A possible solution to cope with the problem of measurement repeatability and to mitigate the impact of the low equivalent resolution of the ADC for the measurement of DC35 and DC37 is to implement averaging. This is a very common solution implemented in many instruments in order to improve the measurement accuracy.

Experiments have been conducted in order to determine the appropriate number of averaging that leads to a sufficient measurement accuracy. Practically, we start from the measurement set that has been collected on IC4 over the different runs, which contains 100 values for each one of the four IMs involved in the model, both for the nominal and degraded configurations. For each configuration, the procedure is then the following:

1. Select k samples among the 100 collected samples (random selection).
2. Compute the mean value of each IM over the k selected samples.
3. Predict power level variation using the mean value of each IM as inputs of the model.
4. Repeat steps 1 to 3 1,000 times to get statistical samples.
5. Repeat steps 1 to 4 for different values of k.

This procedure has been applied for twelve different values of k ranging between 1 (no averaging) up to 100. Results are summarized in Figure 13 that shows the dispersion observed on the values of predicted power level variation over the 1,000 random iterations, for the different values of k (i.e. number of averaging) and the two configurations. The expected range in both configurations is also displayed, which corresponds to the $\pm 3\sigma$ model uncertainty.

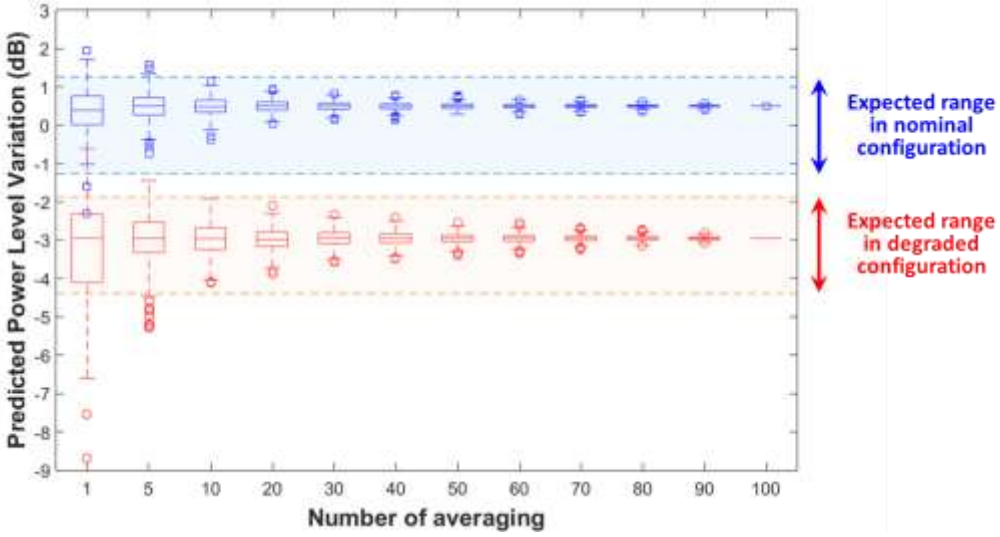

Fig. 13. Boxplot of predicted power level varaition over 1,000 random iterations vs. number of averaging

These results clearly show that, when there is no averaging, the predicated power level variation can be outside the expected range and that there is an overlap between the values predicted in the nominal and degraded configurations, which might comprise the detection of a power level degradation. As expected, the dispersion on the predicted values reduces as the number of averaging increases. With averaging by 5, there is no more overlap between the values predicted in the nominal and degraded configurations but some predicted values are still outside the expected range. With averaging by 10 or higher, all the predicted values remain within the expected range. For the following, we choose to apply averaging by 20.

6.4 Results

The updated code that includes averaging by 20 has been launched on IC4 under five different configurations of the device, i.e. the nominal configuration (C172) and 4 other configurations (C12, C132, C241, C246) spread across the power level variation range observed in the collected dataset.

Results are illustrated in Figure 14, which shows the power level variation predicted from the indirect measurements versus the actual power level variation (measured on the ATE with RF resources), for the five configurations. Two values have been computed in each configuration, using either the original indirect measurements collected on the ATE or the embedded measurements performed within the circuit. Numerical values of the prediction error observed for the five configurations are summarized in Table III.

Looking at these results, it can be observed that the predicted power level variation is in good agreement with the actual one. It exists a difference between the values predicted using the original indirect measurements collected on the ATE and the one measured within the circuit, certainly related to a difference in the measurement accuracy, but the prediction error remains lower than 0.8dB and contained within the model uncertainty in both cases. Overall, these results establish a proof-of-concept of the proposed on-line performance monitoring scheme.

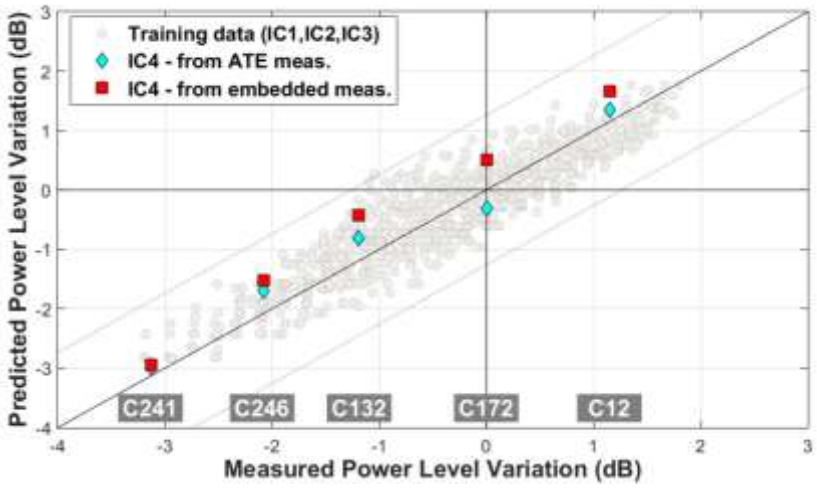

Fig. 14. Predicted power level variation vs. measured power level variation

TABLE III. SUMMARY OF PREDICTION ERRORS ON IC4 IN FIVE DIFFERENT CONFIGURATIONS

	Prediction Error	
	Using ATE meas. & computation on PC	*Using embedded meas. & embedded computation*
C172 (nom.)	-0.31 dB	+0.50 dB
C12	+0.20 dB	+0.52 dB
C132	+0.39 dB	+0.76dB
C241	+0.38 dB	+0.56 dB
C246	+0.15 dB	+0.18 dB

7 Conclusion

We have investigated in this paper the feasibility of implementing on-line RF performance monitoring based on the indirect test strategy. The proposed method implies an initial learning phase in order to build a regression model between the RF performance to be monitored and some indirect measurements available on-chip, from a set of training devices. The second phase takes place during the production of each new manufactured device, in which the model coefficients as well as its own original performance value are stored. The final phase corresponds to the lifetime of the device once deployed in its application: the on-line monitoring process can be launched at any time by performing the embedded measurement of the retained IMs and computation of the performance variation.

A discussion on the type of prediction model and the hardware equipment needed is conducted. The proposed solution was implemented on a practical case of study and results demonstrate the feasibility of on-chip monitoring of the transmitted power level. The indirect measurements available in our experimental vehicle were not specifically tailored towards power level prediction, nor suitable to represent the impact of aging effects, which represents a limitation in this investigation. Further work will address these issues, but a proof of concept is established here.

Funding, Conflicts of Interests, and Competing Interests

Declaration: The authors declare that there is no conflict of interest.

Data availability statement: The data collected and analyzed in this study are not publicly available in order to preserve NXP Semiconductors' industrial confidentiality.

References

[1] M. Nicolaidis, "On-line testing for VLSI," Proc. IEEE International Test Conference, pp. 1042,1997.

[2] H.-G. Stratigopoulos, and Y. Makris, “An adaptive checker for the fully differential analog code,” IEEE Journal of Solid-State Circuits, vol. 41, no. 6, pp. 1421–1429, 2006.

[3] S. K. Devarakond, S. Sen, A. Banerjee, V. Natarajan, and A. Chatterjee, “Built-in performance monitoring of mixed-signal/rf front ends using real-time parameter estimation,” Proc. IEEE International On-Line Testing Symposium (IOLTS), pp. 77–82, 2010.

[4] J. Altet, D. Mateo, and D. Gómez, “On line monitoring of rf power amplifiers with embedded temperature sensors,” Proc. IEEE International On-Line Testing Symposium (IOLTS), pp. 109–113, 2012.

[5] D. Chang, S. Ozev, B. Bakkaloglu, S. Kiaei, E. Afacan, and G. Dundar, “Reliability enhancement using in-field monitoring and recovery for RF circuits,” Proc. IEEE VLSI Test Symposium (VTS), pp. 1–6, 2014.

[6] P. N. Variyam, S. Cherubal, and A. Chatterjee, "Prediction of analog performance parameters using fast transient testing," " IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 21, no. 3, pp. 349-361, 2002.

[7] R. Voorakaranam, S. S. Akbay, S. Bhattacharya, S. Cherubal, and A. Chatterjee, "Signature Testing of Analog and RF Circuits: Algorithms and Methodology," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 54, no. 5, pp. 1018-1031, 2007.

[8] H. Stratigopoulos, “Machine learning applications in IC testing”, Proc. IEEE European Test Symposium (ETS), pp. 1-10, 2018.

[9] L. Abdallah, H. -G. Stratigopoulos, C. Kelma, and S. Mir, “Sensors for built-in alternate RF test,” Proc. IEEE European Test Symposium (ETS), pp. 49-54, 2010.

[10] D. Maliuk, H.-G. Stratigopoulos, H. Huang, and Y. Makris, “Analog neural network design for RF built-in self-test,” Proc. IEEE International Test Conference (ITC), pp. 1–10, 2010.

[11] L. Abdallah, H.-G. Stratigopoulos, and S. Mir, “True non-intrusive sensors for RF built-in test,” Proc. IEEE International Test Conference (ITC), pp. 1–10, 2013.

[12] A. Dimakos, H.-G. Stratigopoulos, A. Siligaris, S. Mir, and E. De Foucauld, “Built-in test of millimeter-wave circuits based on non-intrusive sensors,” Proc. Design, Automation & Test in Europe Conference (DATE), pp. 505–510, 2016.

[13] M. Andraud, H.-G. Stratigopoulos, and E. Simeu, “One-shot non-intrusive calibration against process variations for analog/rf circuits,” IEEE Trans. Circuits & Systems I, vol. 63, no. 11, pp. 2022–2035, 2016.

[14] F. Cilici, G. Leger, M. J. Barragan, S. Mir, E. Lauga-Larroze, and S. Bourdel, “Efficient generation of data sets for one-shot statistical calibration of rf/mm-wave circuits,” Proc. International Conf. on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), pp. 17–20, , 2019.

[15] H. El Badawi, F. Azaïs, S. Bernard, M. Comte, V. Kerzérho, and F. Lefèvre, “Exploring on-line RF performance monitoring based on the indirect test strategy”, Proc. IEEE Latin American Test Symposium (LATS), pp. 1-7, 2021.

[16] B. Halak, “Ageing of Integrated Circuits: Causes, Effects and Mitigation Techniques,” Springer, 2019.

[17] E. Maricau, and G. Gielen, “Analog IC reliability in nanometer CMOS,” Springer, 2013.

[18] P.-I. Mak, and R. P. Martins, “High-/mixed-voltage RF and analog cmos circuits come of age,” IEEE Circuits and Systems Magazine, vol. 10, no. 4, pp. 27–39, 2010.

[19] H. El Badawi, F. Azais, S. Bernard, M. Comte, V. Kerzérho, and F. Lefevre, “Use of ensemble methods for indirect test of RF circuits: can it bring benefits?”, Proc. IEEE Latin American Test Symposium (LATS), pp. 1-6, 2019.

[20] F. Horn, R. Pack, and M. Rieger, “The autofeat python library for automated feature engineering and selection,” Proc. Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 111–120, Springer, 2019.

[21] S. Larguech, F. Azais, S. Bernard, M. Comte, V. Kerzerho, and M. Renovell, “A Generic Methodology for Building Efficient Prediction Models in the Context of Alternate Testing,” Proc. IEEE International Mixed-Signal Test Workshop (IMSTW), pp. 1-.6, 2015.