

ROBUST ESTIMATION OF THE C_p INDEX VIA WAVELETS AND RESAMPLING

ESTIMATIVA ROBUSTA DO ÍNDICE C_p VIA WAVELETS E REAMOSTRAGEM

ESTIMACIÓN ROBUSTA DEL ÍNDICE C_p MEDIANTE ONDÍCULAS Y REMUESTREO

Regis Nunes Vargas¹, Raquel Romes Linhares²

DOI: 10.54899/dcs.v23i92.6020

Recibido: 02/07/2026 | Aceptado: 03/07/2026 | Publicación en línea: 17/07/2026.

ABSTRACT

Statistical Process Control (SPC) is essential for ensuring quality in manufacturing, but it faces challenges in modern production lines due to data corrupted by vibrations and sensor failures. The capability index (C_p), widely used in engineering, is highly sensitive to these outliers, which frequently generates false alarms, unnecessary machine stoppages, and increased production costs. To mitigate this problem, this work presents a robust algorithmic methodology for estimating the C_p index in industrial environments subject to measurement noise and outliers. The proposed approach integrates the Wavelet Shrinkage smoothing technique, capable of attenuating high-frequency noise while preserving signal characteristics, with a Leave-One-Out (Jackknife) resampling logic, using the median as the final estimator to isolate extreme anomalies. Validation was performed through computational experiments simulating a contaminated stable process. Simulated results demonstrate that, while the classical estimator critically fails by classifying a capable process as incapable ($C_p=0.67$) due to sample standard deviation distortion, the proposed algorithm recovers the metric's stability ($C_p=1.22$). By approaching the ideal theoretical value ($C_p=1.33$), the methodology proves its effectiveness in protecting quality indicators, thereby optimizing Production Planning and Control (PPC) decisions.

Keywords: Statistical Control. Wavelet Shrinkage. Jackknife. Capability. Manufacturing.

RESUMO

O Controle Estatístico de Processo (CEP) é essencial para garantir a qualidade na manufatura, mas enfrenta desafios em linhas de produção modernas devido à coleta de dados corrompidos por vibrações e falhas de sensores. O índice de capacidade (C_p), amplamente utilizado na engenharia, é altamente sensível a esses valores atípicos (outliers), o que frequentemente gera falsos alarmes, paradas de máquina desnecessárias e aumento nos custos de produção. Para mitigar este problema, este trabalho apresenta uma metodologia algorítmica robusta para estimar o índice C_p em ambientes industriais sujeitos a ruído de medição e outliers. A abordagem

¹ Ph.D. in Electrical Engineering, Universidade Federal de Uberlândia, Faculdade de Engenharia Elétrica (FEELT-UFU), Uberlândia, Minas Gerais, Brazil. E-mail: regisnv@yahoo.com.br

² Ph.D. in Mathematics, Universidade Federal de Uberlândia, Instituto de Matemática e Estatística (IME-UFU), Uberlândia, Minas Gerais, Brazil. E-mail: raquelromeslinhares@hotmail.com

proposta integra a técnica de suavização Wavelet Shrinkage, capaz de atenuar ruídos de alta frequência preservando as características do sinal, com uma lógica de reamostragem Leave-One-Out (Jackknife), utilizando a mediana como estimador final para isolar anomalias extremas. A validação foi realizada através de experimentos computacionais simulando um processo estável contaminado. Os resultados demonstram que, enquanto o estimador clássico falha criticamente ao classificar um processo capaz como incapaz ($C_p=0.67$) devido à distorção do desvio padrão amostral, o algoritmo proposto recupera a estabilidade da métrica ($C_p=1.22$). Ao se aproximar do valor teórico ideal ($C_p=1.33$), a metodologia comprova sua eficácia em proteger os indicadores de qualidade, otimizando as decisões do Planejamento e Controle da Produção (PCP).

Palavras-chave: Controle Estatístico. Encolhimento de Ondaletas. Jackknife. Capacidade. Manufatura.

RESUMEN

El Control Estadístico de Procesos (CEP) es esencial para garantizar la calidad en la manufactura, pero enfrenta desafíos en las líneas de producción modernas debido a datos corrompidos por vibraciones y fallos en los sensores. El índice de capacidad (C_p), ampliamente utilizado en la ingeniería, es altamente sensible a estos valores atípicos (outliers), lo que frecuentemente genera falsas alarmas, paradas de máquina innecesarias y un aumento en los costos de producción. Para mitigar este problema, este trabajo presenta una metodología algorítmica robusta para estimar el índice C_p en entornos industriales sujetos a ruido de medición y valores atípicos. El enfoque propuesto integra la técnica de suavizado Wavelet Shrinkage, capaz de atenuar ruidos de alta frecuencia preservando las características de la señal, con una lógica de remuestreo Leave-One-Out (Jackknife), utilizando la mediana como estimador final para aislar anomalías extremas. La validación se realizó a través de experimentos computacionales simulando un proceso estable contaminado. Los resultados demuestran que, mientras el estimador clásico falla críticamente al clasificar un proceso capaz como incapaz ($C_p=0.67$) debido a la distorsión de la desviación estándar muestral, el algoritmo propuesto recupera la estabilidad de la métrica ($C_p=1.22$). Al aproximarse al valor teórico ideal ($C_p=1.33$), la metodología demuestra su eficacia para proteger los indicadores de calidad, optimizando las decisiones de Planificación y Control de la Producción (PCP).

Palabras-chave: Control Estadístico. Contracción por Ondículas. Jackknife. Capacidad. Fabricación.



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUCTION

Statistical Process Control (SPC) is a fundamental pillar in production engineering to ensure the quality and stability of industrial operations [4, 5]. Among the utilized metrics, the capability index C_p is widely adopted to quantify a process's ability to meet design specifications [6]. However, the traditional calculation of C_p is highly sensitive to the presence of outliers and high-frequency noise inherent to the shop floor, such as mechanical vibrations or sensing failures [3].

Traditional filtering methods, such as moving averages, tend to distort the true variability of the process. On the other hand, wavelet-based techniques stand out for their ability to remove noise while preserving crucial signal characteristics [1]. Nevertheless, simple filtering does not negate the impact of extreme anomalies on the calculation of the sample standard deviation.

This work proposes an innovative methodology that encapsulates the Wavelet Shrinkage filtering within an exhaustive resampling loop, ensuring that the capability index reflects the real and continuous behavior of the process, shielding decision-making in Production Planning and Control (PPC).

THEORETICAL BACKGROUND

Classical Capability Index (C_p)

The potential capability index C_p evaluates the process dispersion relative to the specification limits tolerated by the customer, defined by:

$$C_p = \frac{USL - LSL}{6\sigma} \quad (1)$$

Where:

USL and LSL represent the upper and lower specification limits, respectively, and σ is the process standard deviation, frequently estimated by the sample standard deviation [4, 6]. The vulnerability of this formulation lies in the fact that a single outlier inflates the value of σ , driving down the C_p index.

Wavelets

Wavelets constitute a mathematical tool capable of representing functions simultaneously in both the localization and scale domains, allowing for the identification of global and local data characteristics. Unlike classical Fourier series expansions, wavelets have local support, making them particularly suitable for analyzing functions with discontinuities, peaks, or regions of high variability.

Multiresolution analysis is built from two fundamental functions: the scaling function, also called the father wavelet, and the wavelet function, frequently called the mother wavelet. Both satisfy fundamental properties of multiresolution analysis.

The scaling function, denoted by $\phi(x)$, is responsible for representing the low-frequency components of the function, describing its global behavior. It has a unitary integral:

$$\int_{-\infty}^{\infty} \phi(x) dx = 1 \quad (2)$$

which justifies its role in representing the global structure of the function and in preserving low-frequency information. The dilated and translated versions of the scaling function are given by:

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k) \quad (3)$$

Where:

j represents the resolution level

and k represents the position of the function in the domain.

In turn, the mother wavelet function, denoted by $\psi(x)$, is responsible for describing the high-frequency components and the local structures of the function. It has a zero mean:

$$\int_{-\infty}^{\infty} \psi(x) dx = 0 \quad (4)$$

a characteristic that makes it particularly well-suited for detecting local variations, discontinuities, and oscillations in the analyzed function. The wavelet functions associated with the different resolution levels are defined by:

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k) \quad (5)$$

A function $f(x) \in L^2(R)$ can then be represented through the wavelet expansion:

$$f(x) = \sum_k c_{J,k} \phi_{J,k}(x) + \sum_{j=J}^{\infty} \sum_k d_{j,k} \psi_{j,k}(x) \quad (6)$$

The coefficients:

$$c_{J,k} = \int f(x) \phi_{J,k}(x) dx \quad (7)$$

are called approximation coefficients and describe the global behavior of the function at resolution level J . On the other hand, the coefficients:

$$d_{j,k} = \int f(x) \psi_{j,k}(x) dx \quad (8)$$

are called detail coefficients and capture the local variations of the function at different resolution scales. Consequently, the detail wavelet coefficients essentially measure local deviations relative to the global behavior captured by the scaling function.

The Wavelet Shrinkage method, introduced by Donoho and Johnstone (1994) [1], acts directly on this representation, consisting of decomposing the signal, applying a threshold to the detail coefficients ($d_{j,k}$) to attenuate the noise, and then reconstructing the signal. This technique is ideal for the industrial environment, as it eliminates spurious variations without excessively smoothing real changes in the physical state of the process.

Robust Statistics and Resampling

The Leave-One-Out (or Jackknife) technique is a resampling method that consists of evaluating the behavior of an estimator by systematically omitting one observation at a time [2].

When combined with the extraction of the median—a measure of central tendency with a high breakdown point—a highly immune estimator to extreme sample contamination is created [3].

METHODOLOGY

The proposed methodology for calculating the robust capability index follows a deterministic flow of five main steps. Given a set of raw data $X = \{x_1, x_2, \dots, x_n\}$ collected from the process, the following is executed:

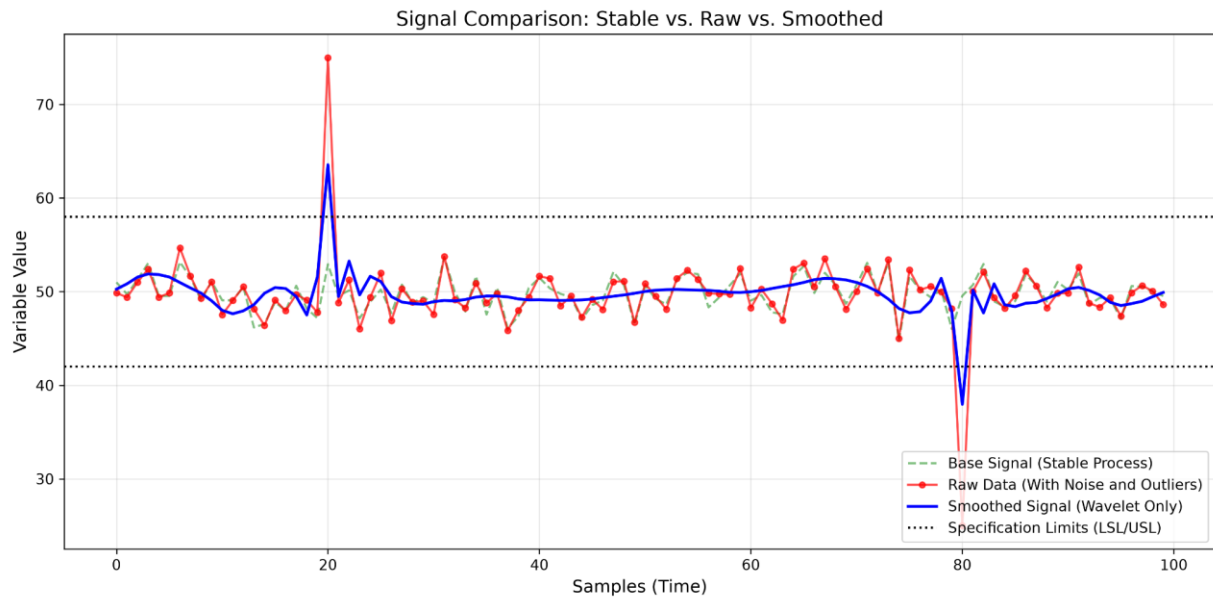
1. For each iteration i from 1 to n , the i -th data point is removed from the sample, generating a reduced subset of size $n - 1$;
2. The Wavelet Shrinkage method is applied to the remaining $n - 1$ data points, producing a smoothed signal free of background noise;
3. The classical capability index (C_p) is calculated for this filtered sample;
4. The calculated index is stored in an array V ;
5. At the end of the n iterations, the median of array V is calculated, which will represent the definitive Robust Capability Index of the process.

Application and Graphical Analysis

To validate the methodology, a computational experiment was conducted simulating a stable industrial process with mean $\mu = 50$ and standard deviation $\sigma = 2$. The specification limits were defined as $LSL = 42$ and $USL = 58$, resulting in an ideal theoretical C_p of 1.33 [6]. The signal was contaminated with normal high-frequency noise and two severe anomalies at samples 20 and 80.

Figure 1 illustrates the comparison among the stable base signal, the contaminated raw data, and the performance of the wavelet filter.

Figure 1. Comparison among the stable signal, the raw data, and the smoothed signal.



Source: Elaborated by the authors (2026).

When applying the traditional metric to the raw data, the calculation was corrupted by the anomalies, returning a $C_p = 0.67$, falsely suggesting an incapable process. The isolated application of the wavelet filter mitigates the noise but does not eliminate the mathematical weight of the outliers [3, 5].

In contrast, the proposed algorithm isolated the extremes during the Leave-One-Out iterations and filtered the residual signal. By extracting the median from the result array, the system neutralized the flawed iterations and returned a C_p of 1.22. The method proved its effectiveness in recovering the true capability underlying the noisy data.

CONCLUSION

This research aimed to answer the following question: how to robustly estimate the process capability index (C_p) in industrial environments contaminated by noise and extreme outliers? The obtained results demonstrated that the integration of Wavelet Shrinkage smoothing, Leave-One-Out resampling, and aggregation by the median provides a mathematically rigorous and highly effective solution for modern Statistical Process Control.

From an academic perspective, this study innovates by proposing an unprecedented combination of multiresolution analysis with robust resampling techniques to isolate anomalies in capability calculations. In practical and social terms, the methodology directly benefits the

manufacturing sector by protecting quality indicators against false alarms generated by noisy sensors or punctual atypical events. This prevents unnecessary machine stoppages and the unwarranted disposal of conforming batches, ultimately optimizing Production Planning and Control (PPC) decisions and promoting more sustainable and resource-efficient industrial operations [4, 5].

Despite the promising results, this research presents some limitations. The proposed algorithm requires a higher computational effort due to the exhaustive resampling loop, and the validation was currently restricted to simulated data based on a normal distribution. For future studies, it is recommended to apply this method to real-time shop floor data and to evaluate its adaptability for other capability indices, such as Cpk, in non-normal processes.

REFERENCES

1. DONOHO, David L.; JOHNSTONE, Iain M. Ideal spatial adaptation by wavelet shrinkage. **Biometrika**, [S.l.], v. 81, n. 3, p. 425-455, 1994.
2. EFRON, Bradley. **The jackknife, the bootstrap and other resampling plans**. [S.l.]: Society for Industrial and Applied Mathematics, 1982.
3. HUBER, Peter J. This is robust statistics. *In: International encyclopedia of statistical science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2025. p. 2783-2786.
4. MONTGOMERY, Douglas C. **Introduction to statistical quality control**. [S.l.]: John Wiley & Sons, 2020.
5. OAKLAND, John S. **Statistical process control**. [S.l.]: Routledge, 2007.
6. PEARN, Wen-lea; KOTZ, Samuel. **Encyclopedia and handbook of process capability indices: a comprehensive exposition of quality control measures**. [S.l.]: World Scientific, 2006.