Accurate estimation of Cpk using an innovative approach based on neural networks

European AEC/APC Conference 2006

Patrice Kiener, *Netral*

<u>François Pasqualini</u>, *STMicroelectronics*

Outline

- Introduction to Capability Index
  - Capability Indices History
  - Why do we have to improve Cpks estimation ?
  - Cpk values associated to fallout
  - Capability Index for the normal distribution
  - General Approach for non normal distribution
  - Quantile Estimation Method
- Neural Network Approach for Distribution Fit
  - Application to Capability estimation
  - Illustrations based on 90nm technology
- Conclusion
- Perspectives

Introduction

- The main idea behind Capability index is to compare what the process is doing with the specification interval

- Capability Indices are becoming today a very valuable decision tool:
  - Internally for Process engineer
    - for qualification and improvement of
      - Tool
      - Process
      - Technology
      - Close loop control (R^2R)
  - Externally for Customers
    - for Quality garantee of their products

- The importance of Capability Indices performance has increased during the 20 last years to reach the quality target equivalent to zero ppm

Capability Indices History

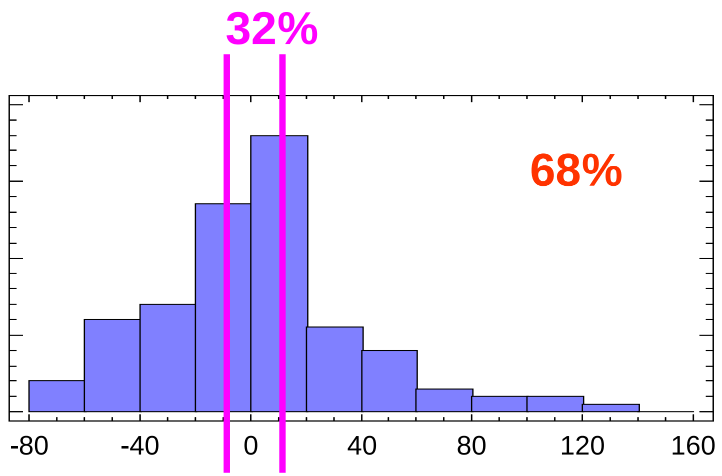
- Statistical Process Control (SPC) is born in 1920st by Walter A. Shewhart
- Capability indices (Cp) were introduced by Juran in 1974
- Ford Motor Company was the first to use aggressively these indices since the early 1980s
- Microelectronics industry has started the use of these index in production in 1986

Today, calculation of Process capabity indices for Key parameters has become a standard in our industry with a very aggressive objective (>1.67)

Why do we have to improve Cpk estimation ?

Parametric Test

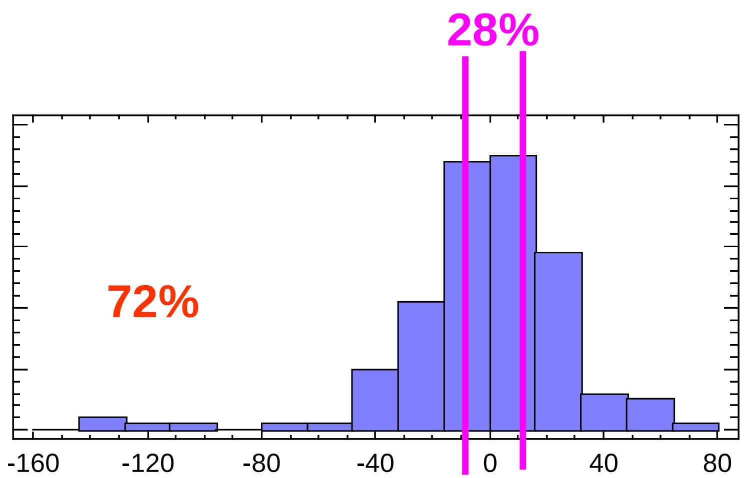
Cpk estimation error using Normal Law (%)



- *Normal law Cpk estimation error is not acceptable for 68% of parameters*

Process

Cpk estimation error using Normal Law (%)



- *Normal law Cpk estimation error is not acceptable for 72% of parameters*

Cpk values associated to fallout

Cpk	One-sided spec.	Two-sided spec.
0.25	226628 ppm	453255 ppm
0.5	66807 ppm	133614 ppm
1	1350 ppm	2700 ppm
1.3	48 ppm	96 ppm
1.67	0.167 ppm	0.334 ppm
2	0.0009 ppm	0.0018 ppm

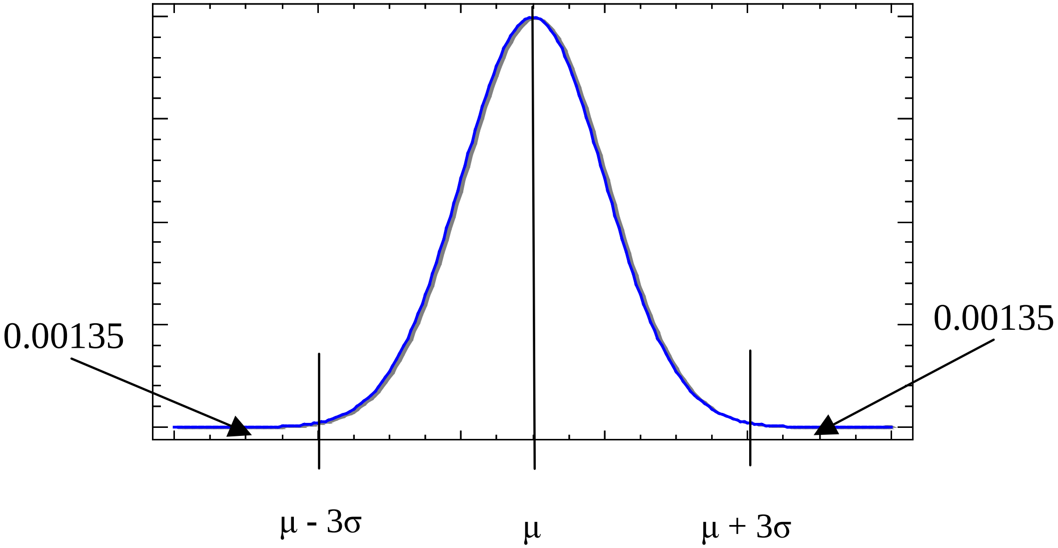
Capability Index for the normal distribution (1/2)

- Definition of Distribution Spread :
  - spread within which almost all of values within a distribution will fall
  - *Originally described (normal law) as within plus or minus three standard deviation (± 3σ) or six standard deviation (6σ)*

- Capability Index can be defined as the ratio:

$$\frac{\text{Specification Interval}}{\text{Process Spread}}$$

Capability Index for the normal distribution (2/2)



99.73 % of data are between μ ± 3σ

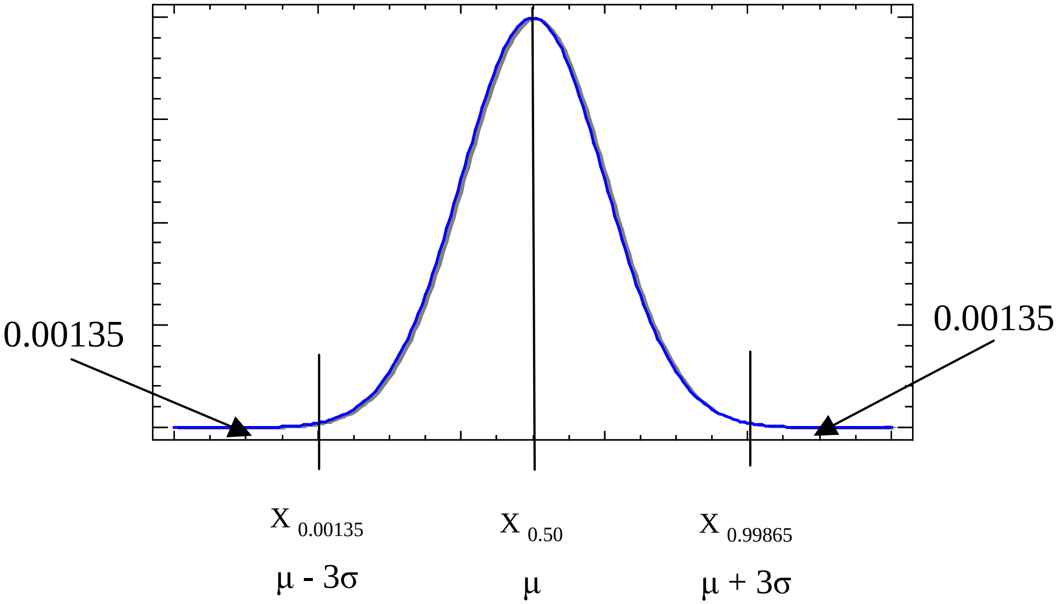
For Normal distribution

- 0.135% of data lies below μ - 3σ
- 50% of data lies below μ
- 99.865 % of data lies below μ + 3σ

$$Cp = \frac{USL - LSL}{6\sigma}$$

$$Cpk = \min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right)$$

General Approach for Normal Distribution



99.73 % of data are between μ ± 3σ or between X $_{0.00135}$ and X $_{0.99865}$

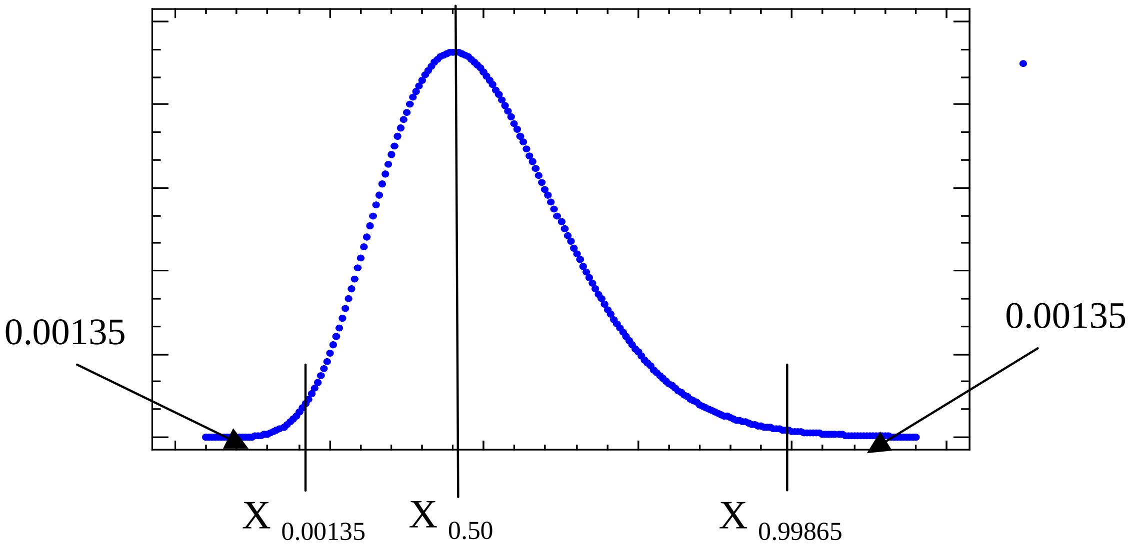
<u>For Normal distribution</u>

- $\mu - 3\sigma = X_{0.00135}$
- $\mu = X_{0.50}$
- $\mu + 3\sigma = X_{0.99865}$

$$Cp = \frac{USL - LSL}{X_{0.99865} - X_{0.00135}}$$

$$Cpk = \min\left(\frac{USL - X_{0.50}}{X_{0.99865} - X_{0.5}}, \frac{X_{0.5} - LSL}{X_{0.50} - X_{0.00135}}\right)$$

General Approach for Non-normal Distribution



99.73 % of data are between
$X_{0.00135}$ and $X_{0.99865}$

For Non-Normal distribution

- 0.135% of data lies below $\mathbf{X}_{0.00135}$
- 50% of data lies below $\mathbf{X}_{0.50}$
- 99.865 % of data lies below $\mathbf{X}_{0.99865}$

$$Cp = \frac{USL - LSL}{X_{0.99865} - X_{0.00135}}$$

$$Cpk = \min\left(\frac{USL - X_{0.50}}{X_{0.99865} - X_{0.5}}, \frac{X_{0.5} - LSL}{X_{0.50} - X_{0.00135}}\right)$$

Quantile Estimation Methods

- Among all existing methods to estimate the three quantiles, the Johnson and Pearson systems of distribution are the best for classical approaches.

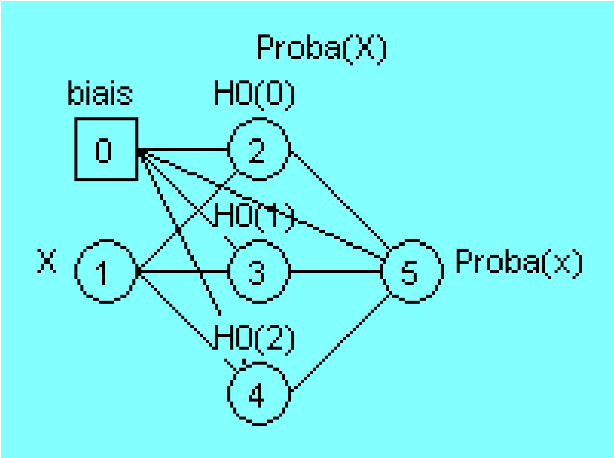
- The Johnson system has been choosen in Crolles in 2001 for its ability to fit distribution with a wide variety of shapes.

- But the remaining difficulties of these methods are:
  - Impossibility to fit some distributions
  - Don’t provide a perfect estimation for some type of distribution as multi-modal distributions
  - Don’t provide confidence interval

A new Approach using Neural Network for Distribution Fit

- Neural Networks are nonlinear functions used for black-box modelling. They have a high flexibility.

- We use them to describe the ordered statistics :
  - One input x_i, one output i / (N+1) in]0..1[
  - 2 or 3 hidden neurons, sometimes 4.
  - Typical function (3 HN) with 10 coefficients :



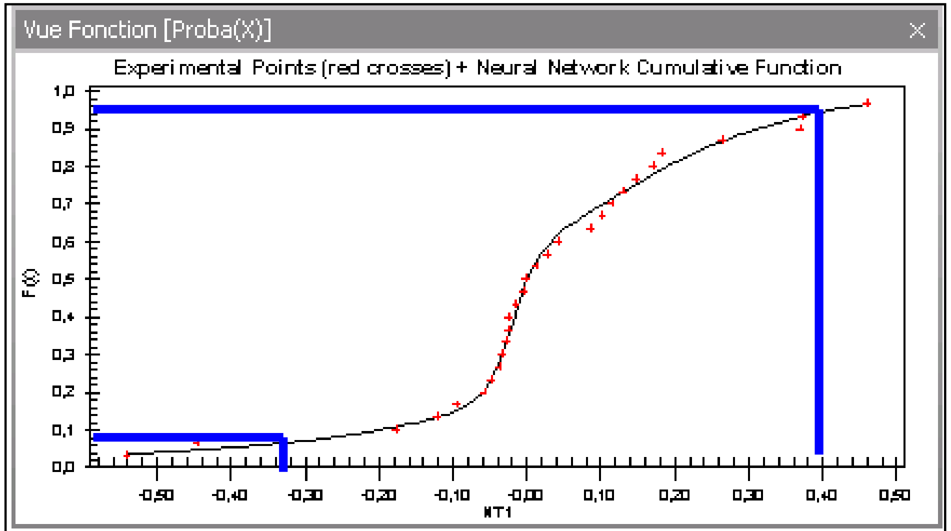
$$y_i = \frac{1}{1+\exp(a_1 + a_2 th(a_3 + a_4 x_i) + a_5 th(a_6 + a_7 x_i) + a_8 th(a_9 + a_{10} x_i))} \cong \frac{i}{N+1}$$

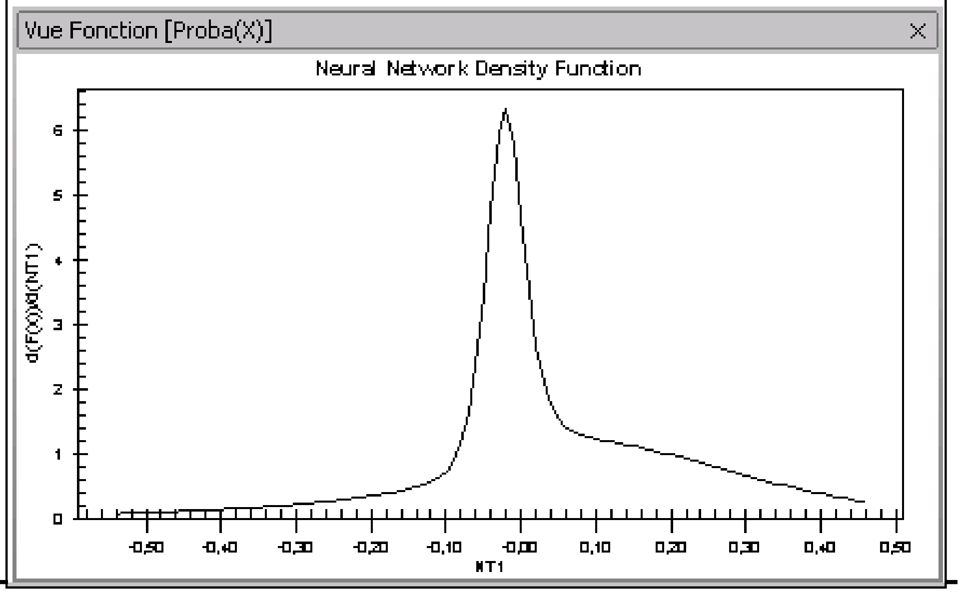
Neural Networks Approach for Distribution Fit

- Training of NN is done on the ordered statistics (red crosses) with a weighted least-square nonlinear regression. It returns an estimate of the Cumulative Function (black line)

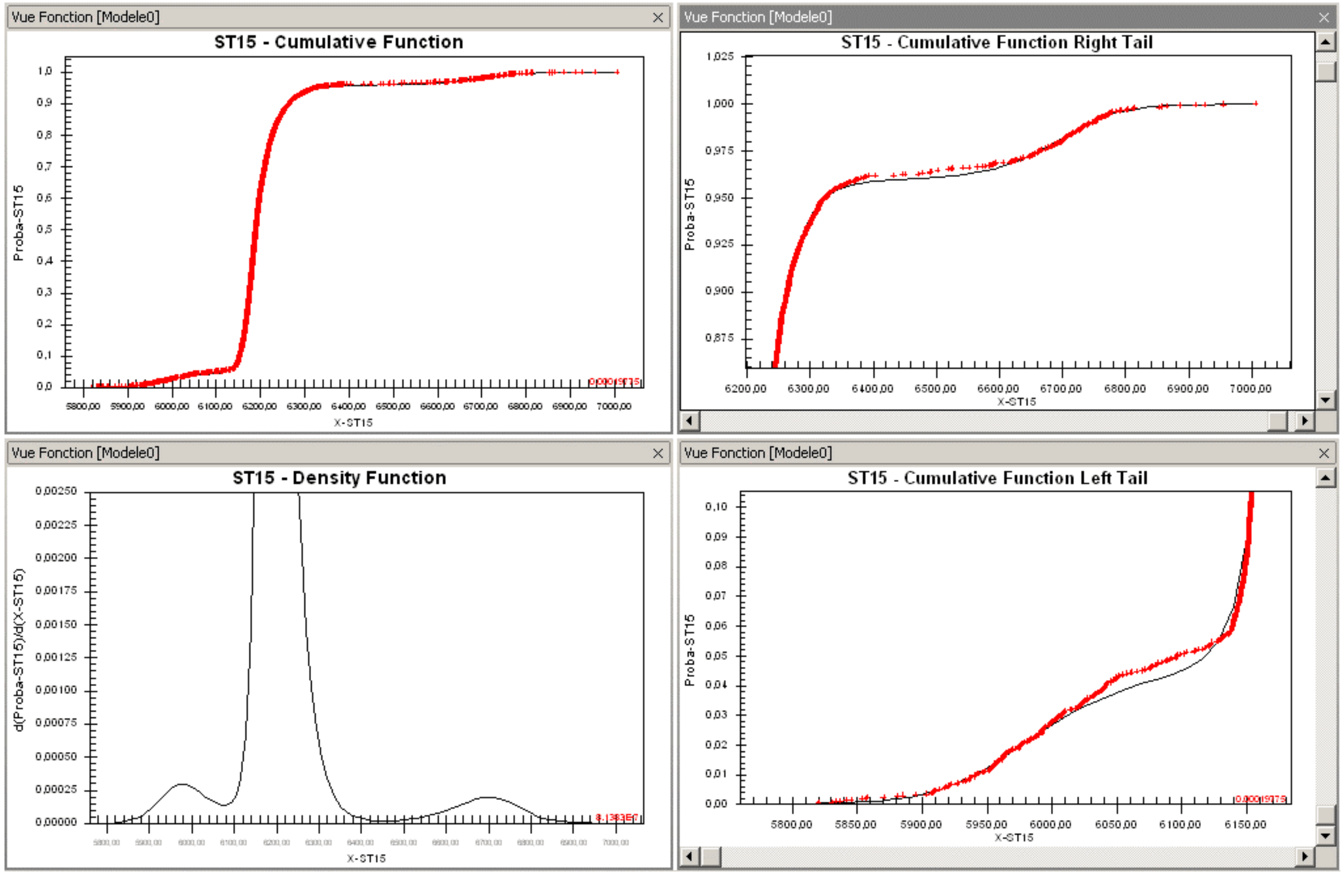
- Quantiles and Cpk are calculated from the Cumulative Function

- By design, NN easily provides the first derivatives of the function. With one single input in the NN, this first derivative is an estimate of the Density Function (black line)





Examples based on 90nm technology

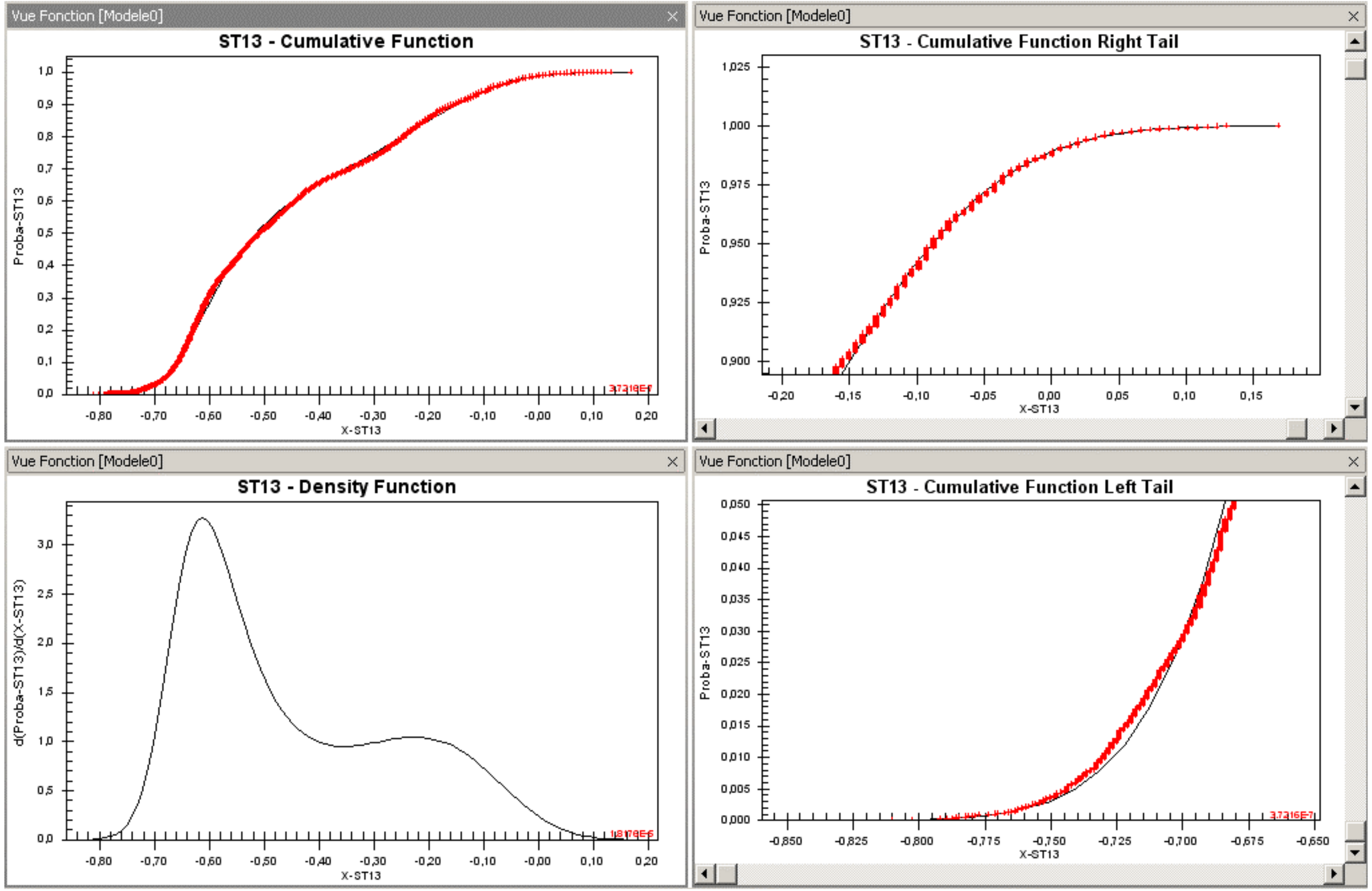


<u>Process: example1</u>

- Cpk-Gaussian = 2.19
- Cpk-NeuralNetwork = 1.30

➢ Overestimation

Examples based on 90nm technology

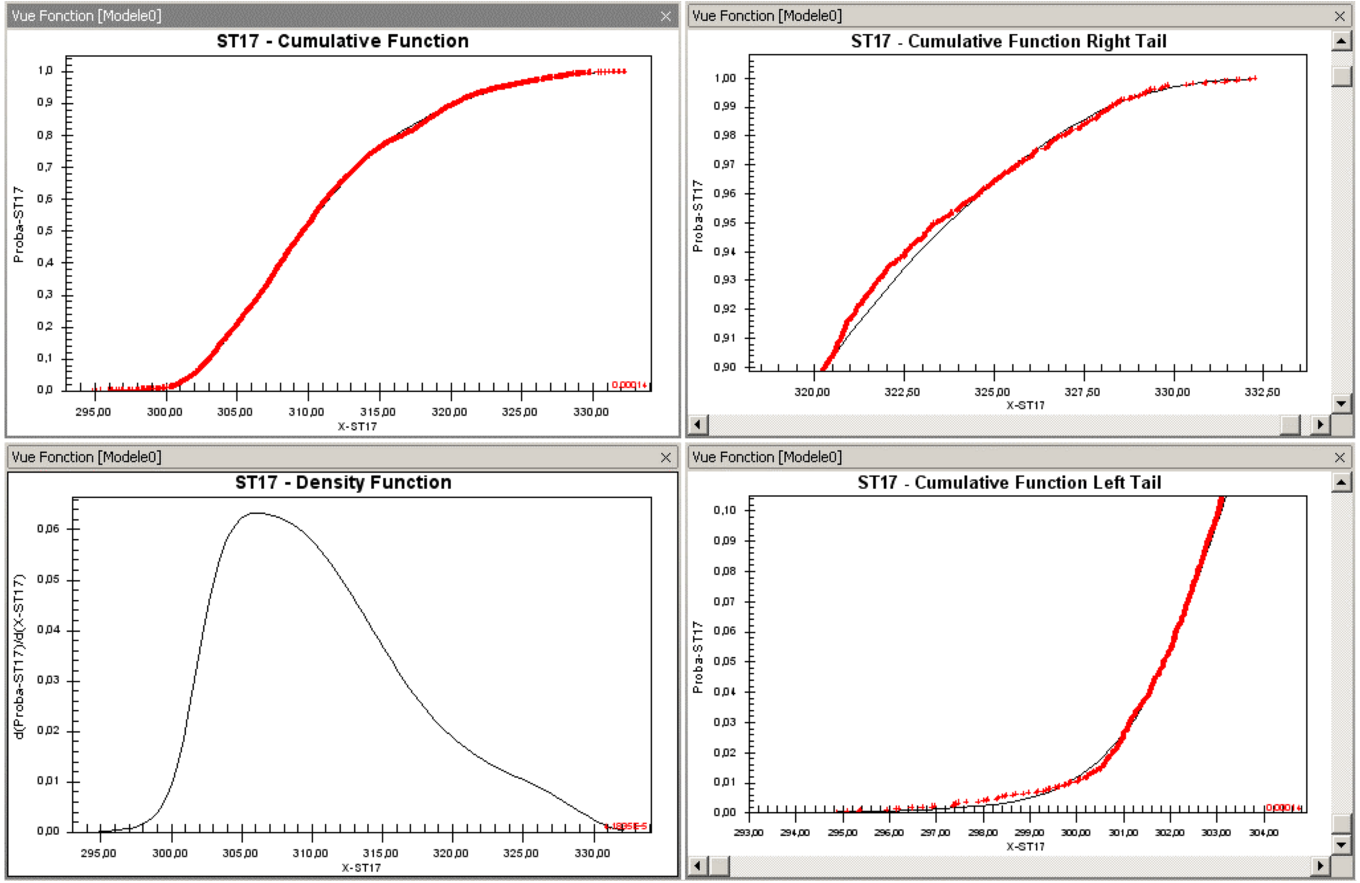


PT : example2

- Cpk-Gaussian = 0.94
- Cpk-NeuralNetwork= 1.93

➢ Underestimation

Examples based on 90nm technology

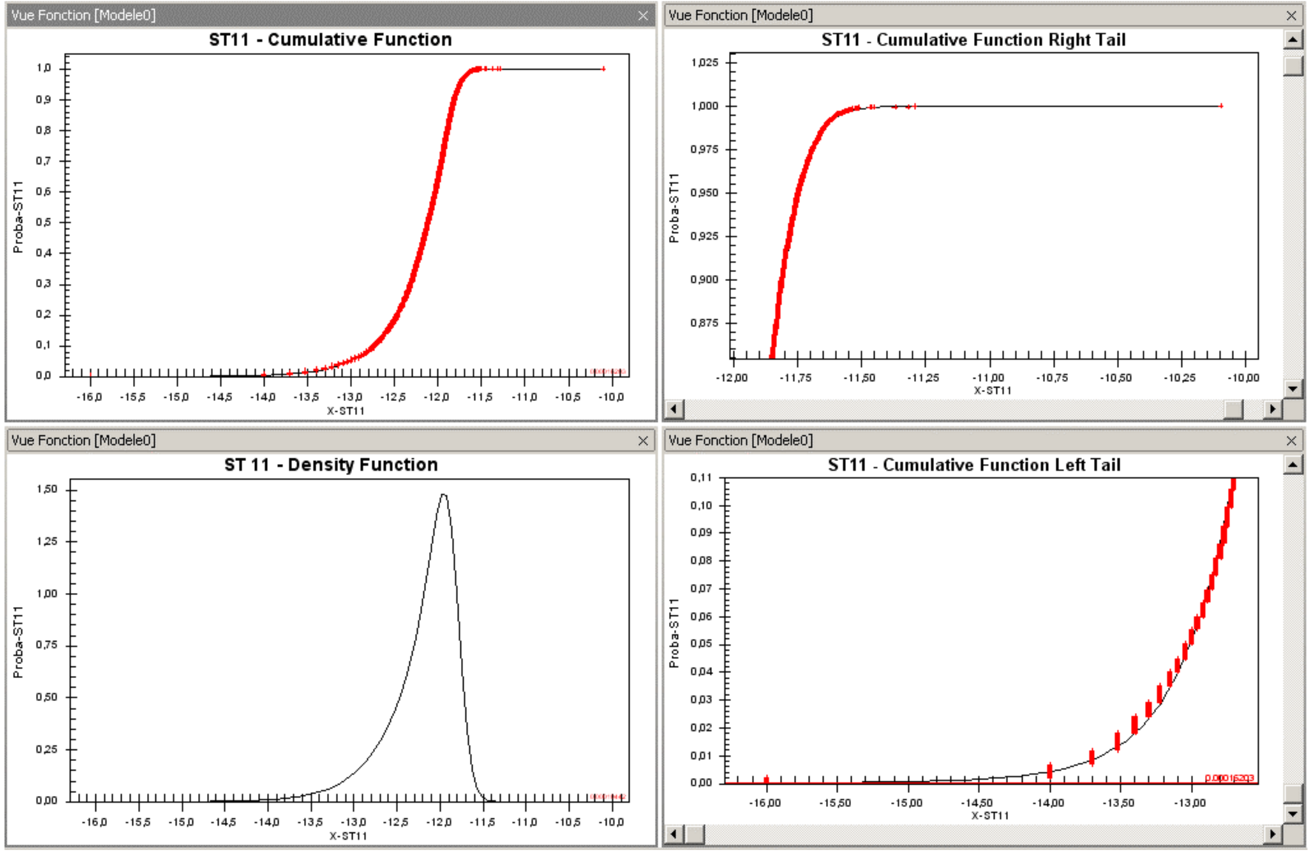


Process: example3

- Cpk-Gaussian = 1.29
- Cpk-NeuralNetwork= 1.73

➢ Underestimation

Examples based on 90nm technology

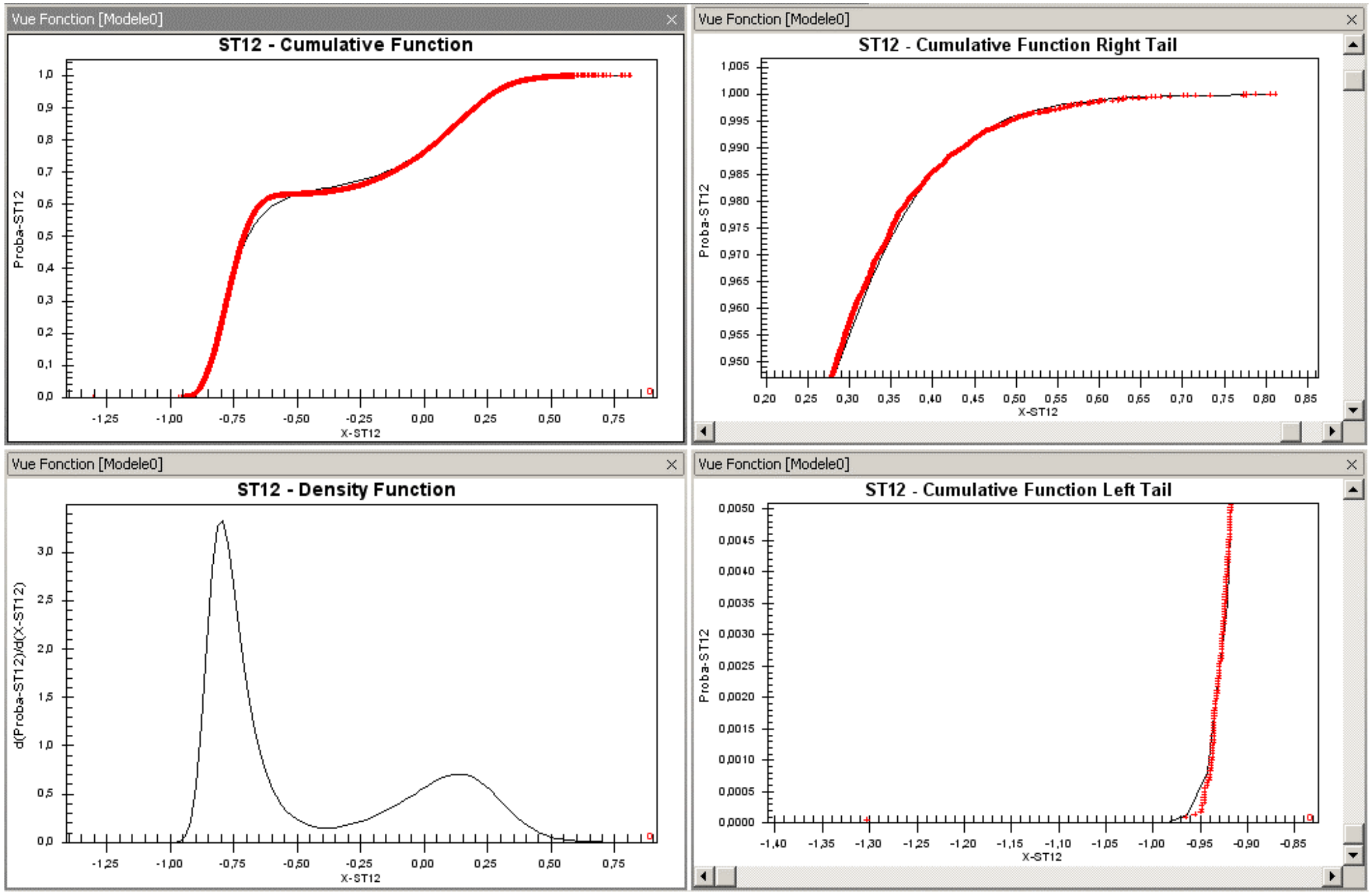


PT : example4

- Cpk-Gaussian = 0.61
- Cpk-NeuralNetwork= 1.18

➢ Underestimation

Examples based on 90nm technology

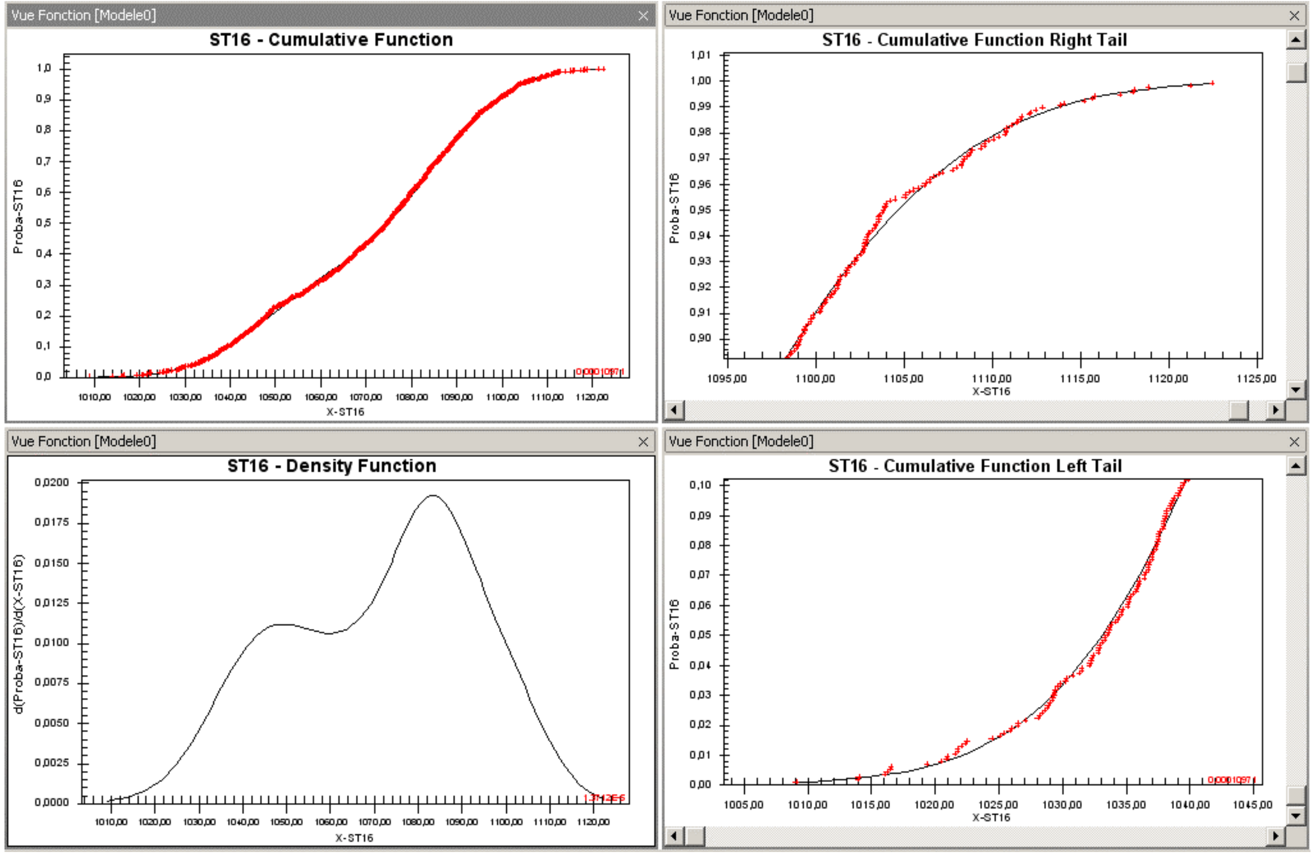


PT : example5

- Cpk-Gaussian = 0.42
- Cpk-NeuralNetwork= 1.26

➢ Underestimation

Examples based on 90nm technology

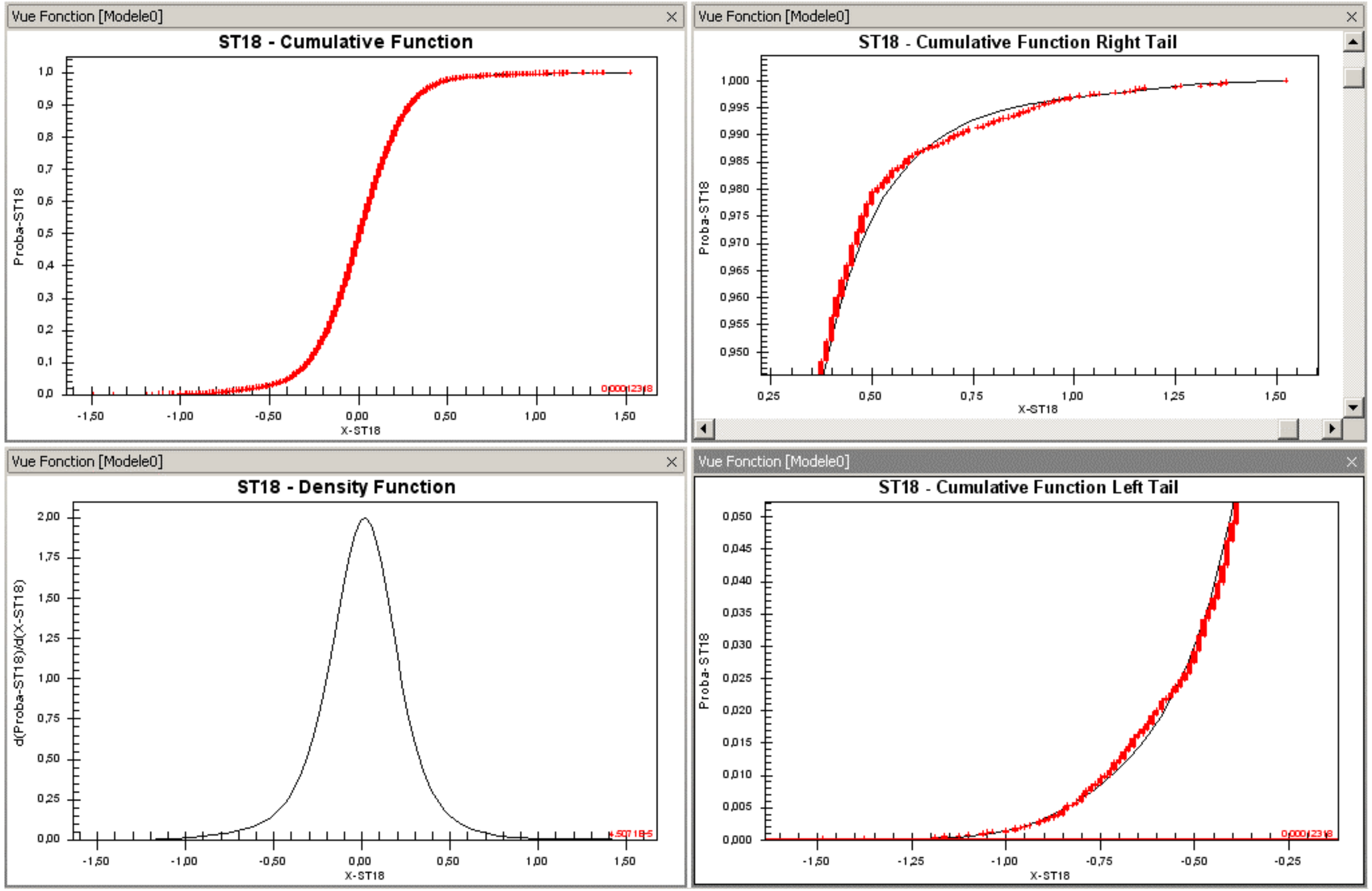


<u>Process: example6</u>

- Cpk-Gaussian=1.17
- Cpk-NeuralNetwork=1.61

➢ Underestimation

Examples based on 90nm technology



Process: example7

- Cpk-Gaussian = 1.30
- Cpk-NeuralNetwork= 0.82

➢ Overestimation

Long tails are the enemy of the Gaussian normal Law !!

Neural Networks Approach for Distribution Fit

- Discrepancy between Neural network approach and Gaussian normal law approach is even better with small set of points
- Estimation starts from 30 points.

Conclusions

Neural Network approach permits to extend the limitation of classical approaches for Quantile estimation

- Independant of shape
- Independant of number of raw data
- Calculation always possible
- Error estimation (on going)

Neural Network approach increases the accuracy of CPK calculation for any distribution type.

Perspectives

- Industrialization of Neural network method for Cpks calculation and performance at Fab level,
- « Bootstrap technique » for robust estimation of small set of points,
- Calculation of Confidence Intervals on Quantile estimation.

This work has been performed in the context of MEDEA+ European Project 2T102

« High Yield MaNufacturing Excellence in sub 65nm CMOS »
(HYMNE)