

Neither Rule Is Safe: Distribution Shape and Mean Shift in Tolerance Stack-Up Analysis

Mukund Kumar Pandey

Assistant Manager (Sales), Essae Digitronics Pvt. Ltd.

Email: maddy.muku2@gmail.com

ABSTRACT

Assembly tolerances are predicted by one of two rules. Worst-case stacking adds component tolerances arithmetically and guarantees assembly but demands component tolerances that are often unmanufacturable. Root-sum-square stacking adds them in quadrature and is far less demanding, at the price of an implied defect rate that is rarely stated. Both rules make assumptions about component distributions that modern supply practice frequently violates. This study quantifies the conservatism of worst-case stacking, the defect rate implied by root-sum-square stacking, and the sensitivity of both to component distribution shape, process mean shift and process capability. Linear stacks of 2 to 20 components, each with a tolerance of ± 0.05 mm, were simulated with 400,000 assemblies per case. Component deviations were drawn as normal, uniform, right-skewed and bimodal distributions, the last representing components from a supplier who has sorted the population and shipped the outer bins. Process mean shifts of 0 to 1.5 standard deviations per component and capability indices of 1.00, 1.33 and 1.67 were examined. The ratio of the worst-case to the root-sum-square prediction grew as the square root of the component count, from 1.41 at two components to 4.47 at twenty. The probability of a stack reaching its worst case was 0.003 % at two components and indistinguishable from zero at four or more, so worst-case stacking guards against an outcome that does not occur. With capable, centred, normally distributed components the root-sum-square prediction was accurate, with 0.27 % of assemblies outside it. Distribution shape broke it: uniform components gave 8.4 % outside and bimodal components 20.9 %. Mean shift was more damaging still, with a 1.0σ shift per component placing 29.1 % of assemblies outside the prediction and a 1.5σ shift 74.9 %. Raising capability from 1.00 to 1.67 reduced the exceedance from 0.27 % to 0.0003 % with no change to the stacking rule. The choice between the two rules matters far less than whether the component distributions satisfy either. Mean shift and supplier sorting invalidate root-sum-square stacking by one to two orders of magnitude, and neither is detected by a tolerance analysis that takes component tolerances as its only input.

Keywords: tolerance analysis; stack-up; root sum of squares; process capability; mean shift; design for manufacture.

1. INTRODUCTION

Every assembly of more than one part requires a prediction of how the variation in its components combines. The prediction determines the tolerances placed on drawings, and those tolerances determine what the parts cost, which process makes them and whether they can be made at all. It is among the most consequential calculations in mechanical design and among the least examined [3,7,16].

Two rules are in general use. Worst-case stacking assumes every component simultaneously sits at the extreme of its tolerance band in the direction that hurts, and adds the tolerances arithmetically [4,5,16]. It guarantees that no assembly can fail, and for stacks of more than a few components it demands tolerances that are expensive or impossible.

Root-sum-square stacking treats component deviations as independent random variables and adds their standard deviations in quadrature [6,9,10]. It is far less demanding and is justified by the observation that components rarely all sit at their extremes at once. It does not guarantee assembly; it implies a defect rate, which is a function of assumptions about the component distributions that the calculation itself does not check.

Those assumptions have become less reliable rather than more. Components are bought rather than made, so the designer does not know the process that produced them or whether it was centred. Suppliers sort and grade, shipping the population that meets a specification while retaining or diverting the rest, which can leave a distribution with the centre removed. Processes drift, and a process running to one side of its tolerance band still passes inspection [8,15].

This study quantifies the consequences. Four questions are addressed: how conservative worst-case stacking actually is; what defect rate root-sum-square stacking implies under its own assumptions; how far distribution shape and mean shift invalidate it; and how these compare with simply requiring more capable components.

2. METHOD

2.1 The Two Rules

For a linear stack of n components each of tolerance $\pm T$, the two predictions for the assembly tolerance are

$$T_{\text{worst case}} = nT, T_{\text{RSS}} = T\sqrt{n} \quad (1)$$

The ratio between them is $\sqrt{n}$, so the rules diverge as the stack grows and agree only for a single component. The root-sum-square form follows from adding variances of independent random variables and assumes, implicitly, that each component tolerance corresponds to the same number of standard deviations and that the resulting sum is approximately normal [5,6].

The relation between a component tolerance and its standard deviation is the process capability index [11,12,14]. A component produced at a capability of 1.00 has a standard deviation of $T/3$, so its tolerance band is $\pm 3\sigma$; at 1.33 it is $\pm 4\sigma$ and at 1.67 it is $\pm 5\sigma$. The same drawing tolerance therefore corresponds to very different amounts of variation depending on who makes the part, and the stacking rule cannot see the difference.

2.2 Simulation

Four hundred thousand assemblies were simulated for each case by summing n independently drawn component deviations. The four distribution shapes used are listed in Table 1 [10], and the simulation parameters in Table 2.

Table 1: Component distributions simulated.

Distribution	Represents	Standard deviation
Normal	A capable, centred process	$T/3$ at a capability of 1.00
Uniform	A process filling its tolerance band, or a coarse adjustment	$T/\sqrt{3}$, larger than the normal case

Distribution	Represents	Standard deviation
Right-skewed	A process bounded on one side, such as a tool-wear-limited dimension	Matched to the normal case
Bimodal	Components from a supplier who has sorted the population and shipped the outer bins	Largest of the four

Table 2: Parameters.

Parameter	Value
Component tolerance, T	± 0.05 mm
Components in the stack, n	2, 4, 6, 10, 20
Assemblies simulated per case	400,000
Process capability, C_p	1.00, 1.33, 1.67
Mean shift per component	0, 0.5, 1.0, 1.5 standard deviations
Assembly requirement used in Section 3.5	± 0.12 mm

Note: All components in a stack are given the same tolerance and the same distribution, which isolates the effect under study. Real stacks mix component types and processes; Section 4.3 discusses what that changes.

3. RESULTS

3.1 How Conservative Is Worst-Case Stacking?

Table 3 and Figure 1 compare the two predictions with the simulated assembly distribution.

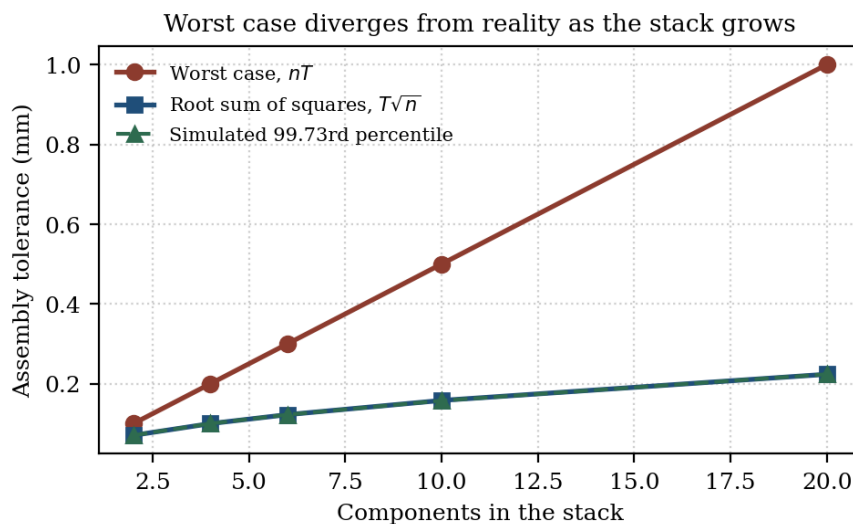


Fig.1: Worst-case and root-sum-square predictions against the simulated 99.73rd percentile.

Two things are visible at once. The root-sum-square prediction tracks the simulated 99.73rd percentile almost exactly at every stack length — 0.1224 against 0.1225 at six components — which is the sense in which it is correct under its own assumptions. And, as Figure 2 shows, the worst case diverges from it as $\sqrt{n}$, reaching a factor of 4.47 at twenty components.

Table 3: The two predictions and what actually happens.

n	Worst case (mm)	RSS (mm)	Ratio	Simulated 99.73rd percentile (mm)	Assemblies beyond worst case
2	0.100	0.0707	1.41	0.0708	0.0033 %
4	0.200	0.1000	2.00	0.1001	< 0.0001 %
6	0.300	0.1225	2.45	0.1224	< 0.0001 %
10	0.500	0.1581	3.16	0.1576	< 0.0001 %
20	1.000	0.2236	4.47	0.2237	< 0.0001 %

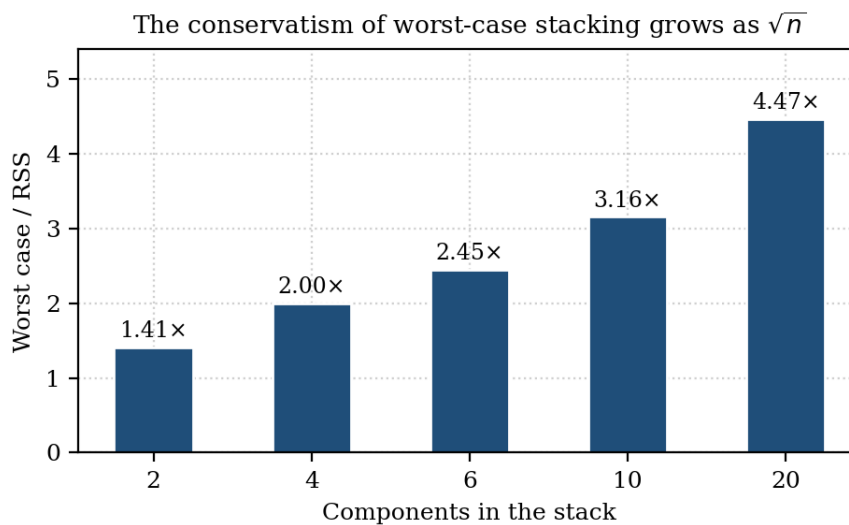


Fig.2: Ratio of the worst-case to the root-sum-square prediction.

The final column of Table 3 states the case against worst-case stacking more sharply than the ratio does. In 400,000 simulated assemblies of four or more capable, centred components, not one reached the worst-case limit. At two components the probability is 0.0033 %, or one assembly in thirty thousand. Designing to a bound that occurs once in never is not conservatism; it is a mis-specification of the problem, and it is paid for in component tolerances that are two to four times tighter than necessary.

3.2 What Root-Sum-Square Stacking Costs

Under its own assumptions the root-sum-square prediction corresponds to the 99.73rd percentile, so 0.27 % of assemblies lie outside it — 2,700 parts per million. That figure is the price of the method and it is very often not stated when the method is chosen. For a product built in tens of thousands it is a meaningful number of rework or scrap events, and whether it is acceptable is a business decision rather than an engineering one.

It is also the best case. Everything that follows examines what happens when the assumptions behind it do not hold.

3.3 Distribution Shape

Table 4 and Figure 3 give the effect of component distribution shape at six components.

Table 4: Effect of component distribution shape at six components.

Distribution	Assembly standard deviation (mm)	Assemblies outside the RSS prediction	Simulated 99.73rd percentile (mm)
Normal	0.0407	0.28 %	0.123
Right-skewed	0.0408	0.58 %	0.139
Uniform	0.0708	8.41 %	0.202
Bimodal (sorted supply)	0.0936	20.92 %	0.249

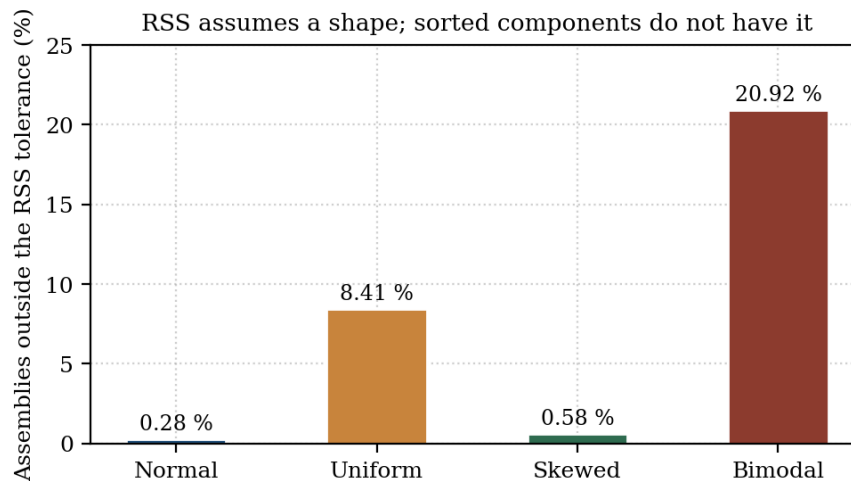


Fig.3: Proportion of assemblies falling outside the root-sum-square prediction, by component distribution.

The normal case behaves as advertised and the skewed case is barely worse, at 0.58 %: the central limit theorem does its work, and six skewed components sum to something close enough to normal that the prediction survives. This is the reassuring half of the result and it is why root-sum-square stacking works as often as it does.

The uniform and bimodal cases are not reassuring. A uniform distribution has a standard deviation of $T/\sqrt{3}$ rather than $T/3$, so components filling their tolerance band carry $\sqrt{3}$ times the variation the rule assumes, and 8.41 % of assemblies fall outside. The bimodal case is worse again at 20.92 %: one assembly in five exceeds a prediction that the analysis reported as covering 99.73 %.

The bimodal case deserves emphasis because it is a consequence of normal commercial practice rather than of a defective process. A supplier who sorts a population and ships the parts that meet a tighter grade elsewhere leaves a distribution with its centre removed. Every part still conforms to the drawing, incoming inspection passes every one of them, and the stack-up analysis — which sees only the tolerance on the drawing — is wrong by a factor of seventy-five in its defect prediction.

3.4 Mean Shift

Table 5 and Figure 4 give the effect of a mean shift in each component at six components.

Table 5: Effect of a mean shift per component at six components.

Shift per component	Assembly mean (mm)	Assembly SD (mm)	Outside RSS prediction	Outside on one side
0	0.000	0.0408	0.27 %	0.13 %
0.5σ	0.050	0.0408	3.82 %	3.81 %
1.0σ	0.100	0.0408	29.08 %	29.08 %
1.5σ	0.150	0.0408	74.89 %	74.89 %

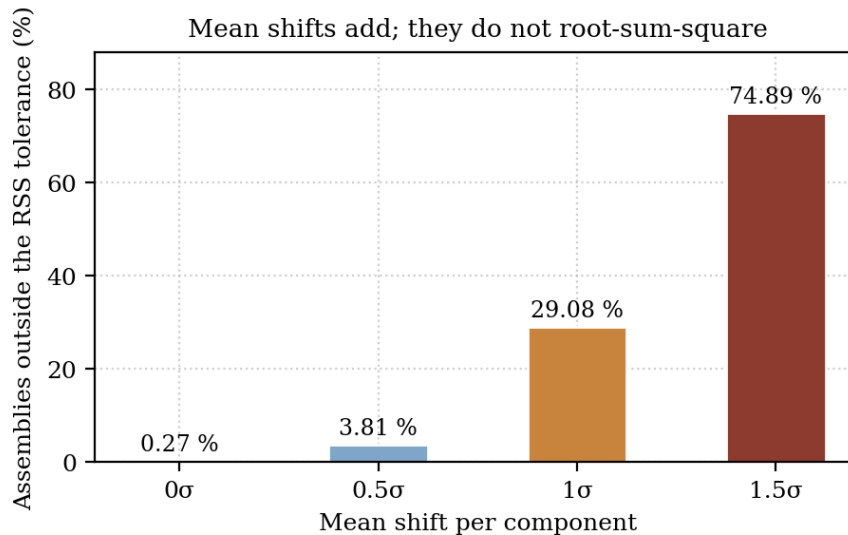


Fig.4: Assemblies outside the root-sum-square prediction against mean shift per component.

This is the most damaging effect examined and the reason is structural. Random deviations combine in quadrature, so n of them grow as $\sqrt{n}$. Mean shifts are not random; they add arithmetically, so n of them grow as n . A shift that is small relative to a single component's variation becomes large relative to the assembly's.

The numbers make the point. A mean shift of one standard deviation per component — a process running slightly off centre but comfortably inside its tolerance band, which would pass any conventional capability check on individual parts — places 29.1 % of assemblies outside the root-sum-square prediction. At 1.5σ , the shift commonly assumed in six-sigma practice as a long-term allowance [13], three quarters of assemblies are outside.

The assembly standard deviation is unchanged at 0.0408 mm in every row. The spread of the assembly is exactly what the rule predicts; it is simply in the wrong place. A tolerance analysis that predicts variation and not location is therefore silent about the dominant failure mechanism, and inspection of component variation alone will not reveal it.

3.5 What the Rules Demand of the Supplier

Table 6 and Figure 5 give the component tolerance each method permits for a fixed assembly requirement of ± 0.12 mm.

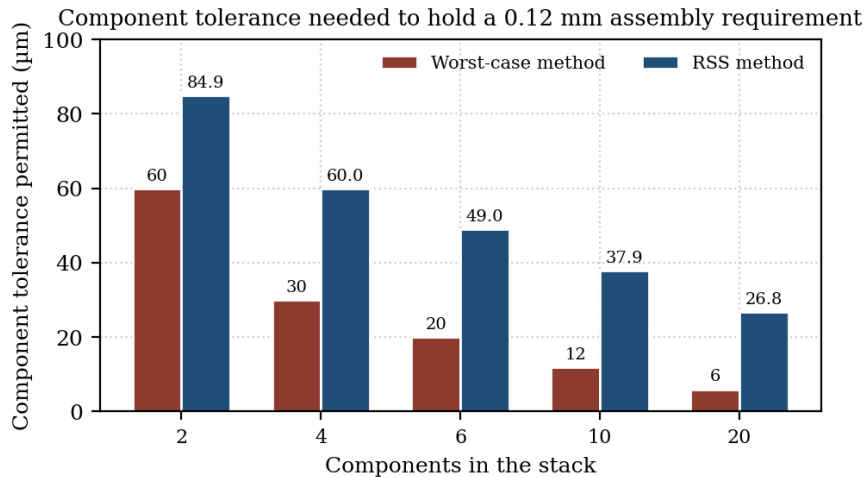


Fig.5: Component tolerance permitted by each method for a fixed assembly requirement.

Read as a manufacturing specification rather than as an arithmetic result, the table explains why root-sum-square stacking is used despite its implied defect rate. At ten components the worst-case rule demands 12 µm where the root-sum-square rule permits 37.9 µm, and the gap is the difference between a ground and a turned feature on most materials.

Table 6 also points at the alternative that neither rule considers. As Figure 6 shows, raising component capability from 1.00 to 1.67 — specifying not just a tolerance but the variation within it — reduced the proportion of assemblies outside the root-sum-square prediction from 0.27 % to 0.0003 %, a factor of nine hundred, with no change to the drawing tolerance or the stacking rule.

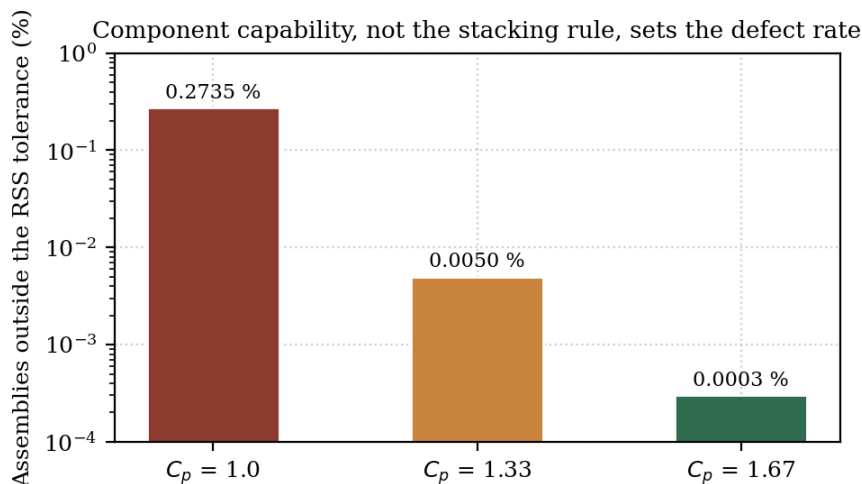


Fig.6: Assemblies outside the root-sum-square prediction against component capability, on a logarithmic scale.

That is a far larger effect than the choice between the two rules, and it is bought in a different currency: not tighter tolerances, which suppliers price steeply, but a capability requirement and the measurement data to demonstrate it [14,17]. For a supplier already holding a capable process it costs nothing but the reporting.

Table 6: Component tolerance permitted to hold a ± 0.12 mm assembly requirement.

n	Worst-case method (μm)	RSS method (μm)	Ratio	Practical comment
2	60.0	84.9	1.41	Both readily achievable
4	30.0	60.0	2.00	Worst case begins to constrain process choice
6	20.0	49.0	2.45	Worst case demands precision machining
10	12.0	37.9	3.16	Worst case approaches grinding tolerances
20	6.0	26.8	4.47	Worst case is impractical for most processes

4. DISCUSSION

4.1 Principal Findings

Four results follow. Worst-case stacking guards against an outcome that did not occur once in 400,000 simulated assemblies of four or more components, while demanding tolerances 2.0 to 4.5 times tighter than the alternative. Root-sum-square stacking is accurate under its assumptions, matching the simulated 99.73rd percentile to three decimal places, and implies a defect rate of 2,700 parts per million that is rarely stated. Distribution shape breaks it by up to a factor of seventy-five, with sorted supply the worst case at 20.9 % outside. And mean shift breaks it furthest, because shifts add while random deviations root-sum-square: 29.1 % outside at a 1.0σ shift per component.

4.2 Implications for Practice

Table 7 summarises each finding, its implication and the corresponding practice.

Table 7: Implications.

Finding	Implication	Practice
Worst case never occurred at $n \geq 4$	The rule protects against nothing	Reserve worst-case stacking for safety-critical or very short stacks
RSS implies 2,700 ppm	The method has a stated price	Quote the implied defect rate whenever RSS is used
Sorted supply gave 20.9 % outside	Conforming parts can invalidate the analysis	Specify distribution or capability, not only a tolerance band
Uniform components gave 8.4 % outside	A tolerance band is not a distribution	Ask what process made the part before assuming a shape
1.0σ shift gave 29.1 % outside	Location error dominates spread error	Require evidence of centring, not only of variation
Assembly SD unchanged under shift	Variation-based checks are blind to it	Monitor assembly mean as well as assembly spread
C_p 1.00 $\rightarrow$ 1.67 improved by 900×	Capability beats the stacking rule	Specify capability; it is cheaper than tolerance for a capable supplier

The third and fifth rows together are the paper's practical message. A tolerance stack-up takes drawing tolerances as its input and produces an assembly prediction, and the two mechanisms that most often invalidate that prediction — sorted or oddly shaped distributions

and off-centre processes — are both invisible in the input. Every part conforms, every drawing is met, and the assembly fails at a rate one to two orders of magnitude above the prediction.

The remedy is not a better stacking rule. It is to specify the quantity the rule actually depends on, which is the distribution of the component dimension rather than the width of the band it must fall inside, and to obtain from the supplier the capability and centring data that demonstrate it. Where that data is unavailable — a catalogue part, a distributor, a new supplier — the root-sum-square prediction should be treated as optimistic by an unknown factor, and that uncertainty is a better reason to use worst-case stacking than the arithmetic of extremes ever was.

4.3 Limitations

The stacks modelled are one-dimensional and linear, with every component carrying the same tolerance and the same distribution. Real stacks combine features made by different processes with different tolerances, and mixing a few poorly behaved components with many well behaved ones produces intermediate results not reported here; the mixed case can be computed with the same method and is the realistic one.

Geometric effects are absent. Real assemblies involve angular contributions, form and orientation controls, datum shift and bonus tolerance from material condition modifiers [1,2], and the sensitivity of the assembly dimension to each component is not uniformly one. A full treatment requires a variation model of the assembly rather than a linear sum [3,8], and the conclusions here apply to the linear contribution within such a model.

Components are treated as statistically independent. Parts made on the same machine in the same batch from the same bar stock are not, and positive correlation moves the assembly behaviour toward the worst-case prediction, which is the one condition under which that rule is defensible. This omission is the main respect in which the analysis is generous to root-sum-square stacking.

The distributions used are idealisations. Real sorted populations retain some central material, real skewed processes vary in their skew, and the 1.5σ shift is a convention rather than a measurement. Absolute percentages are therefore specific to the cases modelled; the orderings and the mechanisms — shifts adding, shapes changing the effective standard deviation, capability dominating — are general.

5. CONCLUSION

In 400,000 simulated assemblies, the worst-case stack-up was never reached with four or more components, while the root-sum-square prediction matched the simulated 99.73rd percentile to three decimal places and implied a defect rate of 2,700 parts per million. On that evidence the long-running argument between the two rules is largely settled in favour of the second, with its price stated.

The argument is also largely beside the point. Components drawn from a sorted supply put 20.9 % of assemblies outside the root-sum-square prediction, and a mean shift of one standard deviation per component put 29.1 % outside, because shifts add arithmetically while random deviations combine in quadrature. Neither mechanism is visible in a drawing tolerance, both pass incoming inspection, and both degrade the prediction by one to two orders of magnitude — far more than the choice of rule.

Meanwhile raising component capability from 1.00 to 1.67 improved the same measure by a factor of nine hundred without touching the tolerance or the rule. The conclusion for practice is that tolerance analysis should specify and verify the distribution of component dimensions, not merely bound them, and that a stack-up performed on drawing tolerances alone

should be reported as what it is: a calculation whose accuracy depends on information the designer has not obtained.

DECLARATIONS

Ethical approval. Not applicable. This is a computational study; no human or animal subjects were involved.

Funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of interest. The author declares no conflicts of interest, financial or non-financial, in relation to the subject matter of this manuscript.

Scope disclaimer. No product, supplier, drawing standard or tolerance analysis package is evaluated. The analysis addresses linear one-dimensional stacks and is not a substitute for a geometric variation study conducted under the applicable dimensioning and tolerancing standard.

Parameter disclaimer. Component tolerances, distributions, shifts and the assembly requirement are representative assumptions. Absolute percentages apply only to the cases modelled; the mechanisms are general, as set out in Section 4.3.

Data and code availability. This is a computational study; no experimental or production data were used. All results were generated by the Monte Carlo simulation described in Section 2, run with a fixed random seed so that every reported value can be reproduced exactly from the parameters listed in Table 2. The code is available from the corresponding author upon reasonable request.

Author's contribution. Mukund Kumar Pandey is the sole author of this work and was responsible for the conception and design of the study, development of the tolerance simulation, analysis and interpretation of the results, preparation of the figures and tables, and drafting, revision and final approval of the manuscript. The author is accountable for all aspects of the work.

REFERENCES

1. American Society of Mechanical Engineers. ASME Y14.5: Dimensioning and Tolerancing. New York: ASME; 2018.
2. American Society of Mechanical Engineers. ASME Y14.41: Digital Product Definition Data Practices. New York: ASME; 2019.
3. Chase KW, Parkinson AR. A survey of research in the application of tolerance analysis to the design of mechanical assemblies. *Res Eng Des.* 1991;3(1):23–37.
4. Chase KW, Greenwood WH. Design issues in mechanical tolerance analysis. *Manuf Rev.* 1988;1(1):50–59.
5. Greenwood WH, Chase KW. A new tolerance analysis method for designers and manufacturers. *J Eng Ind.* 1987;109(2):112–116.
6. Nigam SD, Turner JU. Review of statistical approaches to tolerance analysis. *Comput Aided Des.* 1995;27(1):6–15.
7. Hong YS, Chang TC. A comprehensive review of tolerancing research. *Int J Prod Res.* 2002;40(11):2425–2459.
8. Shen Z, Ameta G, Shah JJ, Davidson JK. A comparative study of tolerance analysis methods. *J Comput Inf Sci Eng.* 2005;5(3):247–256.
9. Bender A. Statistical tolerancing as it relates to quality control and the designer. SAE Technical Paper 680490; 1968.

10. Evans DH. Statistical tolerancing: the state of the art. Part II: methods for estimating moments. *J Qual Technol.* 1975;7(1):1–12.
11. Kane VE. Process capability indices. *J Qual Technol.* 1986;18(1):41–52.
12. Kotz S, Johnson NL. Process capability indices — a review, 1992–2000. *J Qual Technol.* 2002;34(1):2–19.
13. Bothe DR. Statistical reason for the 1.5σ shift. *Qual Eng.* 2002;14(3):479–487.
14. Montgomery DC. *Introduction to Statistical Quality Control.* 8th ed. Hoboken: Wiley; 2019.
15. Creveling CM. *Tolerance Design: A Handbook for Developing Optimal Specifications.* Reading: Addison-Wesley; 1997.
16. Fischer BR. *Mechanical Tolerance Stackup and Analysis.* 2nd ed. Boca Raton: CRC Press; 2011.
17. Taguchi G, Chowdhury S, Wu Y. *Taguchi's Quality Engineering Handbook.* Hoboken: Wiley; 2005.