Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Effect of layer thickness on the radial runout and dimensional accuracy of fused deposition modelling -manufactured polylactic acid spur gears: a Box–Behnken experimental study.

Fatiha Rabbaje

rabbajefatiha@gmail.com

Université Moulay Ismail: Universite Moulay Ismail

Abdelmajid Taleb

Moulay Ismail University National School of Arts and Crafts: Universite Moulay Ismail Ecole Nationale Superieure d'Arts et Metiers

Larbi Lasri

Moulay Ismail University National School of Arts and Crafts: Universite Moulay Ismail Ecole Nationale Superieure d'Arts et Metiers

Research Article

Keywords: fused deposition modelling, polylactic acid spur gear, radial runout, dimension over balls, Box–Behnken design

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11745561/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Effect of layer thickness on the radial runout and dimensional accuracy of fused deposition modelling -manufactured polylactic acid spur gears: a Box–Behnken experimental study.

Fatiha Rabbaje[a,*], Abdelmajid Ait Taleb[a], Larbi Lasri[a]

[a] École Nationale Supérieure d'Arts et Métiers (ENSAM), Moulay Ismail University, Meknes, Morocco

* Corresponding author. E-mail: rabbajefatiha@gmail.com

https://orcid.org/0009-0002-4895-0358

Highlights

- Runout of 45 PLA spur gears from a 15-batch Box–Behnken design was traced to raw dial readings.
- Layer thickness explains 85% of runout variation; speed setting and temperature follow.
- Effects were confirmed at gear and batch level to avoid pseudoreplication.
- A once-per-revolution (eccentricity-type) component carries 99.6% of the runout profile.
- Dimension over balls varies strongly but in two plateaus, so its effects are reported as conditional.

Abstract

Polymer spur gears produced by fused deposition modelling (FDM) are increasingly used in light-duty transmissions, yet process studies rarely report functional gear deviations, and replicate gears printed in the same build are often analysed as independent runs. This study quantifies how layer thickness (0.1–0.3 mm), nozzle temperature (190–220 °C) and speed setting (100–140 mm/s) affect the radial runout F_r and the dimension over balls M of polylactic acid (PLA) spur gears (module 2.5 mm, 20 teeth). A Box–Behnken design with 15 batches of three gears (45 gears) was printed; F_r was computed from 20 archived dial-indicator readings per gear and M was read twice per gear over 4.45 mm balls. Every analysis was run at gear level and on the 15 batch means, the experimental unit. Batch configurations explained 97% of the runout variance (within-batch SD 0.019 mm). Layer thickness accounted for 85.2% of the variation, speed setting for 9.2% and temperature for 1.6%; a layer × speed interaction was detected at both levels. A reduced quadratic model reached R^2 =

0.974 without lack of fit, and a first-order model fixed a priori predicted the mean runout of unseen batches with an RMSE of 0.025 mm. Fourier analysis showed that a once-per-revolution component explains a median 99.6% of each runout profile, i.e. runout is essentially an eccentricity of the toothing. The deviation $\Delta M = M - M_0$ varied strongly between configurations (96% of variance, repeatability 0.010 mm) but grouped into two plateaus that a polynomial does not describe; because a measurement-session effect cannot be excluded, its parameter effects are reported as conditional. All gears exceeded the coarsest runout tolerance of ANSI/AGMA 2015-2-A06 (grade C12, 0.190 mm), which only the predicted optimum setting approaches. Thin layers (0.1–0.2 mm) are recommended where runout matters.

Keywords: fused deposition modelling; polylactic acid spur gear; radial runout; dimension over balls; Box–Behnken design.

1. Introduction

Fused deposition modelling (FDM), also called fused filament fabrication, is the most widespread additive manufacturing process for thermoplastics because of its low cost and wide material choice [1–3]. Beyond prototypes, it is now used to make functional parts, among which polymer spur gears for light-duty transmissions, test rigs and spare parts [4–6]. For a gear, geometric accuracy is not a cosmetic attribute: deviations of the toothing position and size govern backlash, transmission error, noise and load sharing.

The influence of FDM process parameters on dimensional accuracy has been studied extensively, mostly on prismatic specimens [7–10]. Layer thickness, extrusion temperature and printing speed are repeatedly identified as influential [11–14], with the stair-step effect, bead geometry and thermal shrinkage as the usual explanations [15, 16]. Studies on printed gears are fewer. They compare materials and printing directions [4, 5], quantify interchangeability and forming precision [6], analyse shrinkage-related deviations with tooth-span, pin and chordal measurements [17, 18], or apply designed experiments to profile deviations [19] and dimensional responses [20, 21]. In the sample of gear studies we examined, none reports the radial runout, although runout is a standard functional indicator of the toothing position relative to the axis [22], and the dimension over balls or pins, the usual workshop check of tooth size, is reported only for single process settings [17].

Two methodological issues also limit the strength of the conclusions that can be drawn from such experiments. First, several gears are often printed together in one build for each

configuration; these gears are subsamples of one experimental unit, and treating them as independent runs inflates the degrees of freedom of the error term, a form of pseudoreplication [23, 24]. Second, a dimension over balls determines only one geometric quantity; converting it into a "pitch diameter" requires an assumption on the form of the deviation, which is seldom stated.

This work addresses these gaps on PLA spur gears produced according to a Box–Behnken design [25, 26]. Its contributions are: (i) a runout dataset traceable to the raw dial readings of 45 gears, with an explicit data-integrity check; (ii) an analysis of variance conducted at gear level and at batch level, so that each conclusion is labelled with its strength of evidence; (iii) a harmonic decomposition showing what the measured runout physically represents; (iv) a treatment of the dimension over balls directly as a measured deviation ΔM, with its repeatability, its sensitivity to the ball diameter and the conditions under which its parameter effects can be trusted. Section 2 describes the materials, metrology and statistics, Section 3 the results, and Section 4 discusses their interpretation and limitations.

2. Materials and methods

2.1. Gear, material and printing

The gear is an unmodified (zero profile shift) spur gear with module m = 2.5 mm, z = 20 teeth and pressure angle α = 20°, giving a reference diameter d = 50.000 mm, a base diameter d_b = $d \cos \alpha$ = 46.985 mm, a tip diameter of 55 mm and a root diameter of 43.75 mm. Face width is 10 mm and the 20 mm bore carries a keyway. A single parametric CAD model (CATIA V5) was used for all gears. Gears were printed from white PLA filament of 1.75 mm diameter (SUNLU), all from the same spool lot, on a Bambu Lab A1 printer with a 0.4 mm nozzle, sliced in Bambu Studio. Non-studied parameters were held constant: gear axis perpendicular to the build plate (tooth profile in the printing plane), 100% infill, 100% part cooling and a bed temperature of 60 °C. The three gears of a configuration, labelled a, b and c, were placed side by side and printed in one build, hereafter called a batch (Fig. 1a).

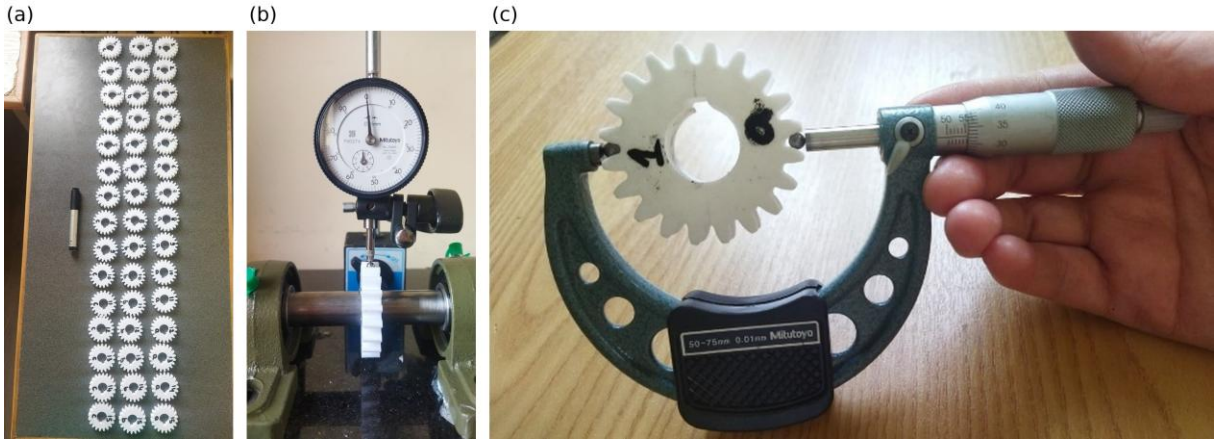


Fig. 1. Experimental campaign and metrology: (a) the 45 gears (15 batches of three); (b) runout measurement, with the gear keyed on a reference shaft supported by two bearings and a ball placed in each tooth space under a dial indicator; (c) measurement of the dimension over balls with an outside micrometer (50–75 mm, 0.01 mm).

2.2. Experimental design

Three factors were studied at three levels (Table 1): layer thickness e, nozzle temperature T and speed setting V. The coded variables are

$$X_1 = \frac{e - 0.2}{0.1}, \quad X_2 = \frac{T - 205}{15}, \quad X_3 = \frac{V - 120}{20} \tag{1}$$

with e in mm, T in °C and V in mm/s. The Box–Behnken design comprises 12 edge-midpoint configurations and three centre configurations printed in separate batches, i.e. 15 batches and 45 gears, printed in the run order given in Table 2. It supports a full quadratic model without combining the extreme levels of all factors [25, 27]. The speed parameter changed in the slicer (all moves or only some) was not recorded, and the build times logged by the slicer were identical for the three speeds at a given layer thickness; V is therefore interpreted as a speed *setting* whose effect on the actual deposition speed is not established.

Table 1. Factors and levels (coded −1, 0, +1) and fixed printing parameters.

Factor	Symbol	−1	0	+1
Layer thickness (mm)	e (X_1)	0.1	0.2	0.3
Nozzle temperature (°C)	T (X_2)	190	205	220
Speed setting (mm/s)	V (X_3)	100	120	140

Fixed: PLA Ø1.75 mm (single spool lot); 0.4 mm nozzle; bed 60 °C; infill 100%; part cooling 100%; gear axis vertical; three gears per build.

Table 2. Box–Behnken design, printing order and batch results (mean ± SD over the three gears of each batch).

Batch	e (mm)	T (°C)	V (mm/s)	Run order	F_r (mm)	ΔM (mm)
1	0.1	190	120	12	0.267 ± 0.023	−0.584 ± 0.018
2	0.1	220	120	6	0.220 ± 0.010	−0.266 ± 0.043
3	0.3	190	120	3	0.567 ± 0.029	−0.631 ± 0.023
4	0.3	220	120	11	0.543 ± 0.015	−0.457 ± 0.053
5	0.1	205	100	8	0.240 ± 0.026	−0.206 ± 0.023
6	0.1	205	140	5	0.283 ± 0.021	−0.221 ± 0.023
7	0.3	205	100	10	0.470 ± 0.017	−0.439 ± 0.036

8	0.3	205	140	1	0.607 ± 0.006	−0.551 ± 0.053
9	0.2	190	100	14	0.393 ± 0.038	−0.601 ± 0.008
10	0.2	190	140	9	0.490 ± 0.010	−0.271 ± 0.028
11	0.2	220	100	13	0.340 ± 0.010	−0.252 ± 0.033
12	0.2	220	140	4	0.450 ± 0.010	−0.276 ± 0.025
13	0.2	205	120	2	0.393 ± 0.012	−0.237 ± 0.028
14	0.2	205	120	15	0.417 ± 0.015	−0.242 ± 0.021
15	0.2	205	120	7	0.407 ± 0.021	−0.259 ± 0.033

$\Delta M = M - M_0$, with M the mean of the two readings of a gear and $M_0 = 56.414$ mm (balls of 4.45 mm).

2.3. Radial runout

Runout was measured on the fixture of Fig. 1b. The gear was keyed and clamped on a reference shaft carried by two bearings; a ball of nominal diameter 4.45 mm was placed successively in each of the 20 tooth spaces and its radial position r_k read with a dial indicator graduated in 0.01 mm. Following the definition of runout in ISO 1328-2 and ANSI/AGMA 2015-2-A06 (Annex B) [22, 28], the probe contacts both flanks at approximately mid tooth depth and F_r is the range of the readings. With 4.45 mm balls, the ball centre lies at a radius of 25.98 mm and the ball–flank contact, located on the line of action at $r_b \tan \alpha_M - d_w/2 = 8.87$ mm from the base-circle tangency point, at a radius of 25.11 mm, i.e. 0.42 mm above mid-depth (24.69 mm):

$$F_r = \max_k r_k - \min_k r_k, \quad k = 1, \ldots, 20 \tag{2}$$

The 20 readings of every gear were archived. F_r is invariant to a common offset of the readings, so the zero setting does not affect it; however, the angular position of tooth space 1 was not referenced identically for all gears, which precludes any analysis of the direction of eccentricity. Before analysis, the recorded maxima, minima and ranges were checked against the 20 readings of each gear. For gear 1c, a first transcription gave 0.60 mm in tooth space 18, inconsistent with its recorded range of 0.28 mm; the raw reading sheet showed 0.06 mm, a transcription error that was corrected. All analyses use the 45 gears, and their robustness to the exclusion of gear 1c is reported.

2.4. Dimension over balls

The dimension over balls M was read with an outside micrometer (Mitutoyo, range 50–75 mm, resolution 0.01 mm) across two balls of diameter $d_w = 4.45$ mm placed in diametrically opposite tooth spaces (Fig. 1c). For an even number of teeth, zero profile shift and a nominal tooth thickness $s = \pi m/2$ (so that $s/d = \pi/2z$), the pressure angle α_M at the ball centre satisfies

$$\operatorname{inv} \alpha_M = \tan \alpha_M - \alpha_M = \operatorname{inv} \alpha + \frac{d_w}{d_b} - \frac{\pi}{2z} \tag{3}$$

which was solved by Newton–Raphson iteration, giving $\alpha_M = 25.286°$. The theoretical dimension over balls is then

$$M_0 = \frac{d_b}{\cos \alpha_M} + d_w = 56.414 \text{ mm} \quad (4)$$

Each gear was read twice (M_1, M_2) on two pairs of tooth spaces, and the analysed response is the deviation

$$\Delta M = \frac{M_1 + M_2}{2} - M_0 \quad (5)$$

ΔM is a direct measurement. Converting it into a pitch-diameter deviation requires an assumption on the form of the deviation. Under a homothetic assumption (uniform scaling of the toothing with balls of fixed size), differentiation of Eqs. (3)–(4) gives, near the nominal point,

$$\frac{dM}{dD_p} = \frac{\cos \alpha}{\cos \alpha_M}\left(1 - \frac{d_w}{d_b \tan \alpha_M}\right) = 0.831 \quad \Rightarrow \quad \Delta D_{p,eq} \approx 1.20\, \Delta M \quad (6)$$

whereas attributing the whole deviation to tooth thickness would give $\Delta M \approx 2.20\, \Delta s$ from the same reading. Because the equivalent pitch diameter depends on such an assumption, all analyses below are made on ΔM, and Eq. (6) is used only to give an order of magnitude. The absolute value of ΔM also depends on the ball diameter, since

$$\frac{\partial M_0}{\partial d_w} = 1 + \frac{1}{\sin \alpha_M} \approx 3.34 \quad (7)$$

so that an error of 0.01 mm on d_w shifts M_0, and every ΔM, by about 0.033 mm without changing comparisons between configurations. The balls were not certified. Measurements were made on several days with the same pair of balls; the allocation of gears to days was not recorded.

2.5. Statistical analysis

Two-level analysis. The experimental unit is the batch; the three gears of a batch are subsamples of one build. Every analysis was therefore run twice: on the 45 gears, and on the 15 batch means, where pure error is estimated from the three centre batches only (2 degrees of freedom). The variance between configurations was first decomposed with a one-way random-effects model,

$$y_{lj} = \mu + b_l + \varepsilon_{lj}, \; b_l \sim N(0, \sigma_b^2), \; \varepsilon_{lj} \sim N(0, \sigma_w^2) \quad (8)$$

where l = 1, …, 15 is the batch and j = 1, 2, 3 the gear. At the centre point, the variance between the three batch means was compared with the value expected from within-batch variation alone, as a test of a build-to-build effect.

Response surface. Each response was described by the full quadratic model in coded variables

$$Y = \beta_0 + \sum_i \beta_i X_i + \sum_{i<j} \beta_{ij} X_i X_j + \sum_i \beta_{ii} X_i^2 + \varepsilon \tag{9}$$

fitted by least squares, with type II analysis of variance at the 5% level [24]. The model was reduced by backward elimination (removal threshold $p > 0.10$, hierarchy respected). A term significant at 5% at both levels is called *detected*; significant at one level only, or with $0.05 < p < 0.10$, a *trend*; otherwise *not detected*, which does not establish absence of effect. Adequacy was assessed with R^2, adjusted R^2 and a lack-of-fit test against pure error. Residuals were checked for normality (Shapiro–Wilk), constant variance (Breusch–Pagan, Levene), independence with respect to run order (Durbin–Watson and Spearman correlation on ordered batches) and influence (studentised residuals, Cook's distance).

Prediction of unseen configurations. The PRESS-based predicted R^2 leaves out one gear at a time while the other gears of its batch remain in the fit, and is therefore optimistic. Predictive ability was instead assessed by leave-one-batch-out validation: each batch was removed in turn, the model fitted on the 14 others and the batch mean predicted,

$$\mathrm{RMSE}_{\mathrm{batch}} = \sqrt{\frac{1}{15}\sum_{l=1}^{15}\left(\overline{y}_l - \hat{y}_{(-l)}\right)^2} \tag{10}$$

Only models whose form is fixed a priori (first-order polynomial and full quadratic model) were evaluated in this way, so that no model choice used the left-out batch [29].

Harmonic analysis. The 20 readings of each gear were decomposed into a Fourier series over one revolution,

$$r_k = a_0 + \sum_{h=1}^{10}\left(a_h \cos\frac{2\pi hk}{20} + b_h \sin\frac{2\pi hk}{20}\right), \quad R_1^2 = \frac{10\left(a_1^2 + b_1^2\right)}{\sum_k (r_k - \overline{r})^2} \tag{11}$$

where R_1^2 is the share of the profile variance explained by the once-per-revolution component, which describes an eccentricity of the toothing with respect to the axis of rotation; its peak-to-peak amplitude is $2(a_1^2 + b_1^2)^{1/2}$. Computations used Python 3.13 with pandas, statsmodels 0.15.0 and SciPy 1.18.1.

3. Results

3.1. Variability of runout

Batch-mean runout ranged from 0.220 to 0.607 mm, whereas the within-batch standard deviation was only 0.019 mm (Fig. 2, Table 2). The random-effects model attributed 97.4% of the runout variance to differences between configurations ($F(14, 30) = 112.5$, $p < 0.0001$). At the centre point, the standard deviation of the three batch means (0.012 mm) matched the value expected without a build effect (0.011 mm; $F(2, 30) = 1.08$, $p = 0.35$): no build-to-build variability was detected, although with two degrees of freedom this test has low power, which is why every conclusion was also checked on batch means.

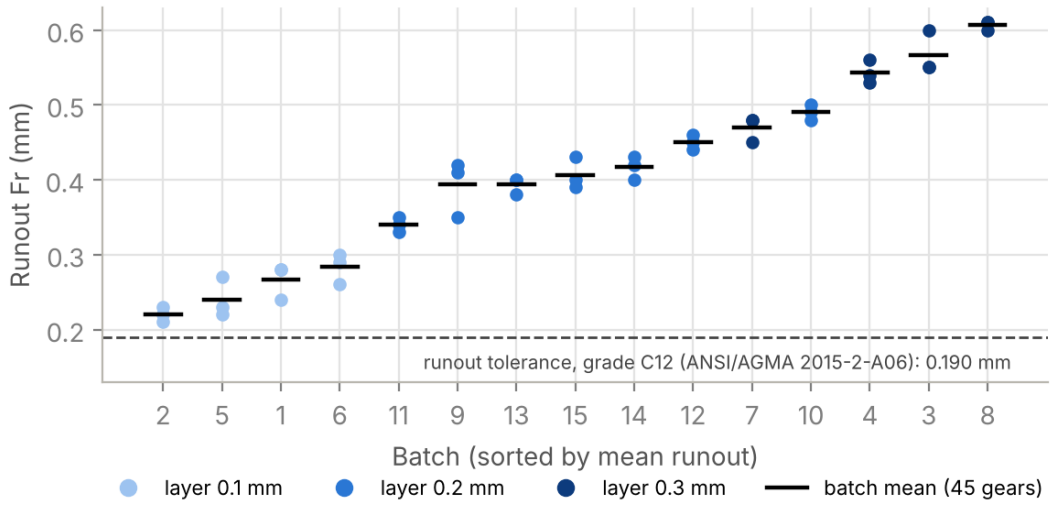


Fig. 2. Runout of the 45 gears (points) and batch means (bars), batches sorted by mean runout; colour indicates the layer thickness. Dashed line: runout tolerance of accuracy grade C12, ANSI/AGMA 2015-2-A06 (0.190 mm).

3.2. Effects of the process parameters

In the full quadratic model fitted to the 45 gears, the three main effects and the X_1X_3 interaction were significant; X_1^2 was borderline ($p = 0.051$) and the other terms were not (X_1X_2: $p = 0.33$; X_2X_3: $p = 0.57$; X_2^2: $p = 0.34$; X_3^2: $p = 0.27$). The hierarchy is pronounced: layer thickness accounts for 85.2% of the total variation, speed setting for 9.2%, temperature for 1.6% and the interaction for 1.1%. Marginal batch means rise from 0.252 mm at 0.1 mm layers to 0.413 mm at 0.2 mm and 0.547 mm at 0.3 mm. Backward elimination retained five terms (Table 3).

Table 3. Analysis of variance of the full quadratic runout model (45 gears; 35 residual df; $R^2 = 0.976$).

Term	Coefficient (mm)	SS (mm²)	*F*	*p*	Share (%)
X_1 (layer)	+0.1471	0.51920	1250.6	< 0.0001	85.2
X_3 (speed setting)	+0.0483	0.05607	135.1	< 0.0001	9.2
X_2 (temperature)	−0.0204	0.01000	24.1	< 0.0001	1.6
X_1X_3	+0.0233	0.00653	15.7	0.0003	1.1
X_1^2	−0.0124	0.00169	4.08	0.051	0.3
X_3^2	+0.0068	0.00051	1.24	0.27	0.1
X_1X_2	+0.0058	0.00041	0.98	0.33	0.1
X_2^2	+0.0060	0.00040	0.95	0.34	0.1
X_2X_3	+0.0033	0.00013	0.32	0.57	0.0
Residual	—	0.01453			2.4

Removing the four non-significant terms raises the residual degrees of freedom from 35 to 39 and changes the residual mean square, so that X_1^2 becomes significant in the reduced model ($p = 0.034$, Table 4) while it is borderline in the full model. With F_r in mm, the reduced model is

$$\hat{F}_r = 0.4129 + 0.1471X_1 - 0.0204X_2 + 0.0483X_3 + 0.0233X_1X_3 - 0.0133X_1^2 \quad (12)$$

Table 4. Analysis of variance of the reduced runout model (45 gears).

Source	df	SS (mm²)	*F*	*p*	Share (%)
X_1 (layer)	1	0.51920	1272.2	< 0.0001	85.2
X_3 (speed setting)	1	0.05607	137.4	< 0.0001	9.2
X_2 (temperature)	1	0.01000	24.5	< 0.0001	1.6
X_1X_3	1	0.00653	16.0	0.0003	1.1
X_1^2	1	0.00197	4.84	0.034	0.3
Residual	39	0.01592			2.6
Lack of fit	7	0.00370	1.38	0.25	
Pure error	32	0.01222			

Table 5. Coefficients of the reduced runout model (coded variables, mm) at the two levels of analysis.

Term	Coefficient	SE	95% CI	*p* (45 gears)	*p* (15 batches)	Status
Intercept	0.4129	0.0044	[0.4039, 0.4218]	< 0.0001	< 0.0001	—
X_1	0.1471	0.0041	[0.1387, 0.1554]	< 0.0001	< 0.0001	detected
X_2	−0.0204	0.0041	[−0.0288, −0.0121]	< 0.0001	0.0016	detected
X_3	0.0483	0.0041	[0.0400, 0.0567]	< 0.0001	< 0.0001	detected
X_1X_3	0.0233	0.0058	[0.0115, 0.0351]	0.0003	0.0057	detected
X_1^2	−0.0133	0.0060	[−0.0255, −0.0011]	0.034	0.079	trend

The model explains 97.4% of the variance of the 45 gears (adjusted $R^2 = 0.971$) and 99.2% of the variance of the 15 batch means (adjusted $R^2 = 0.988$). Lack of fit is not significant at either level ($p = 0.25$ on gears, $p = 0.51$ on batches), and the residual standard deviation (0.020 mm) is close to the within-batch value (0.019 mm). Within the studied domain, increasing the layer thickness from 0.1 to 0.3 mm raises runout by about 0.29 mm, whereas raising the temperature from 190 to 220 °C lowers it by 0.041 mm, about 10% of the mean runout: the temperature effect is detected but of moderate practical importance.

The effect of the speed setting depends on layer thickness (Fig. 3). The four batches crossing these two factors at 205 °C show that, between 100 and 140 mm/s, mean runout increases by 0.043 mm at 0.1 mm layers (batches 5 and 6) and by 0.137 mm at 0.3 mm layers (batches 7 and 8), an effect about three times larger; the reduced model gives 0.050 and 0.143 mm. The interaction is detected at both levels but rests on the four batches that inform it directly.

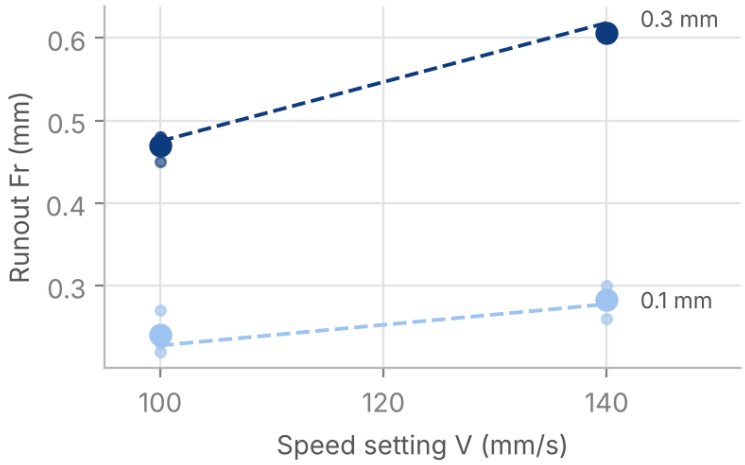


Fig. 3. Layer thickness × speed setting interaction at T = 205 °C: gears (small points), means of batches 5–8 (large points) and reduced-model prediction (dashed).

Response maps of the reduced model over the layer × speed plane, and their use to locate a predicted optimum, are not developed here; only the prediction of the reduced model at the most favourable vertex of the domain is discussed below, in relation to accuracy grades. No such map would be meaningful for ΔM, whose quadratic model shows significant lack of fit (Section 3.5).

The residual diagnostics do not contradict the model assumptions (Supplementary Fig. S1): Breusch–Pagan p = 0.27, Levene across configurations p = 0.99, Shapiro–Wilk p = 0.64 on gears and 0.86 on batches; no drift with printing order was detected (Durbin–Watson on batches 1.76; Spearman ρ = 0.33, p = 0.22). The largest studentised residual is 2.52 and the largest Cook's distance 0.25 (both gear 5c). With 15 batches these tests have limited power, and the measurement order, distinct from the printing order, was not recorded.

3.3. Predictive check and robustness

The PRESS-based predicted R^2 was 0.965, but this statistic does not measure the ability to predict a new configuration. In leave-one-batch-out validation, the first-order polynomial and the full quadratic model predicted the mean runout of an unseen batch with RMSE of 0.025 and 0.029 mm (out-of-sample R^2 0.955 and 0.935), against 0.124 mm for the grand mean.

Neither reaches the floor set by within-batch scatter, $0.019/\sqrt{3} \approx 0.011$ mm. The reduced model, whose terms were chosen on all data, gives 0.018 mm, an optimistic value because the left-out batch took part in term selection. The full quadratic model, which estimates ten coefficients from fourteen batches, does not predict better than the first-order polynomial. Comparing machine-learning models under the same batch-wise protocol is a natural extension of this work.

Re-running the whole analysis without gear 1c, including backward elimination, gave the same retained terms, coefficients equal to within 0.001 mm and the same conclusions at both levels; the X_1X_3 interaction remained detected ($p = 0.0003$ on gears, 0.0057 on batches) and X_1^2 remained a trend (Supplementary Fig. S2).

3.4. Harmonic content of the runout profiles

The once-per-revolution component explained a median 99.6% of the profile variance, and more than 95% for 42 of the 45 gears; its peak-to-peak amplitude averaged 95% of F_r (Fig. 4). The three exceptions are gears 1a (46%) and 1c (9%), among the least eccentric gears, where local defects become relatively visible, and gear 3a (86%). The measured runout is thus almost entirely an eccentricity of the toothing with respect to the axis of rotation, not a sum of local tooth-to-tooth irregularities, consistent with the typical runout diagram of ANSI/AGMA 2015-2-A06, in which eccentricity is a component of runout [28]. Whether this eccentricity belongs to the printed gear (offset between toothing and bore) or partly to the fixture (shaft runout, bore deformation on mounting) cannot be decided with the present data, because the shaft was not measured alone and tooth space 1 was not referenced angularly.

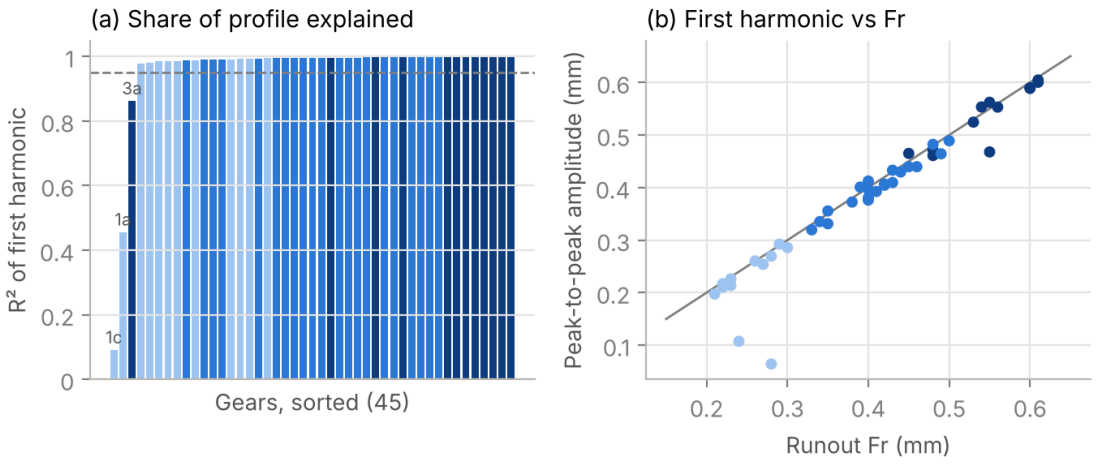


Fig. 4. Harmonic analysis of the 45 runout profiles: (a) share of profile variance explained by the first harmonic (gears sorted); (b) peak-to-peak amplitude of the first harmonic against F_r.

3.5. Dimension over balls

The two readings of each gear give the measurement repeatability: $M_2 - M_1$ has a standard deviation of 0.0145 mm, i.e. a repeatability standard deviation $s_r \approx 0.010$ mm, of the order of the micrometer resolution (Fig. 5b). M_2 exceeded M_1 by 0.010 mm on average (paired t-test, $p < 0.001$), a small systematic offset linked to the second pair of tooth spaces or to the reading order, cancelled in the comparisons because both readings are averaged for every gear.

Across the 45 gears, ΔM had a mean of −0.366 mm and a standard deviation of 0.157 mm (range −0.654 to −0.184 mm). Variability was again dominated by configurations: within-batch SD 0.032 mm, between-configuration share 96% of the variance ($F(14, 30) = 72.7$, $p < 0.0001$). The three centre batches were highly reproducible (−0.237, −0.242 and −0.259 mm; $F(2, 30) = 0.37$, $p = 0.69$). Layer thickness and temperature effects were detected at both levels, with marked curvature along both factors (Table 6). Marginal batch means were −0.319, −0.305 and −0.519 mm for 0.1, 0.2 and 0.3 mm layers, and −0.522, −0.308 and −0.313 mm for 190, 205 and 220 °C.

Table 6. Tests on the deviation of dimension over balls ΔM.

Analysis	Result
One-way random effects (45 gears)	$F(14, 30) = 72.7$, $p < 0.0001$; between-configuration share 96%
Centre batches (build-to-build effect)	$F(2, 30) = 0.37$, $p = 0.69$
Full quadratic model, 45 gears	$R^2 = 0.86$; X_1, X_2, X_2X_3, X_1^2, X_2^2: $p < 0.0001$; X_3: $p = 0.10$
Full quadratic model, 15 batches	$R^2 = 0.89$; X_1: $p = 0.025$; X_2: $p = 0.021$; X_1^2: $p = 0.048$; X_2^2: $p = 0.053$
Lack of fit (reduced model)	gears: $F(5, 32) = 25.6$, $p < 0.0001$; batches: $F(5, 2) = 65.9$, $p = 0.015$
Leave-one-batch-out RMSE	grand mean 0.164 mm; first-order 0.151 mm; full quadratic 0.205 mm

The quadratic surface describes these data poorly: lack of fit is significant at both levels and, in leave-one-batch-out validation, no polynomial clearly beats the grand mean. The batches indeed fall into two distinct groups rather than along a smooth surface (Fig. 5a). Nine batches, including the three centre batches, have ΔM close to −0.24 mm (−0.21 to −0.28 mm), whereas six lie between −0.44 and −0.63 mm: the three batches printed at 190 °C with 100 or 120 mm/s (batches 1, 3, 9) and the three 0.3 mm batches printed at 205 or 220 °C (batches 4, 7, 8). Batch 10 (190 °C, 140 mm/s) belongs to the first group, which produces the apparent temperature × speed interaction. Residuals ordered by printing sequence show positive autocorrelation (Durbin–Watson 1.02), although no monotonic trend with order was found (Spearman $\rho = -0.06$).

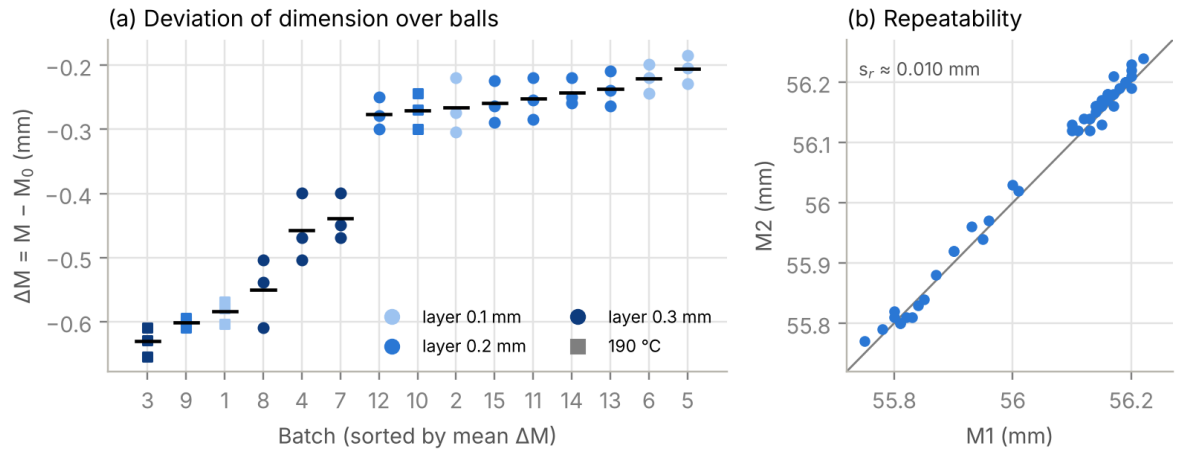


Fig. 5. (a) Deviation of dimension over balls ΔM for the 45 gears (points) and batch means (bars), batches sorted by mean ΔM; squares: 190 °C. (b) Repeatability: second reading M_2 against the first M_1.

Runout and ΔM were negatively correlated across gears ($r = -0.45$, $p = 0.002$); the correlation of batch means was similar but not significant ($r = -0.46$, $p = 0.08$). This association mainly reflects the common effect of layer thickness, which increases runout and decreases ΔM; measuring over diametrically opposite tooth spaces cancels a pure eccentricity to first order, so the two responses carry distinct information.

4. Discussion

4.1. What the runout results establish

Table 7 summarises the strength of evidence for the main findings. The dominance of layer thickness on runout is the most robust result: it is detected at both levels, unchanged without gear 1c, and large compared with the within-batch scatter. The measurements do not give access to the underlying physics, and the interpretations below are hypotheses consistent with the direction of the effects. A thicker layer increases the stair-step effect and the bead cross-section, hence the amount of material that solidifies and shrinks at each pass [15, 16]. Thermo-mechanical simulations of printed gears show localised shrinkage and warpage driven by deposition thermal gradients [18]. The harmonic analysis suggests that, in the present gears, such phenomena translate into a global offset between toothing and axis rather than into local form defects; whether the offset arises between layers, from asymmetric contraction of the bore or rim, or from the fixture cannot be identified with the available measurements.

The layer × speed interaction has a parsimonious explanation: at constant extrusion width, volumetric flow is proportional to the product of layer thickness and speed, and a common

effect of both on one quantity produces a positive interaction. This hypothesis requires that the speed setting changes the actual deposition speed, which the identical build times at 100, 120 and 140 mm/s do not confirm; it should be tested by identifying the slicer parameter and comparing flow rates with the hot-end limit. The favourable effect of temperature is consistent with the lower viscosity of molten PLA, which eases flow and inter-bead bonding [13, 15].

Table 7. Strength of evidence of the main conclusions.

Conclusion	Evidence	Status
Layer thickness dominates runout	85% of variation; detected at both levels; unchanged without gear 1c	Established
Runout increases with speed setting and decreases moderately with temperature	Detected at both levels; temperature effect ≈ 10% of mean runout	Established (speed as a setting)
Layer × speed interaction	p = 0.0003 (gears), 0.0057 (batches); four informative batches	Detected; needs confirmation
Runout profile is an eccentricity	Median 99.6% of variance in first harmonic	Established; origin (gear or fixture) open
Effect hierarchy predicts unseen configurations	Leave-one-batch-out RMSE 0.025 mm (first-order)	Established within the domain
Layer and temperature change ΔM (up to ≈ 0.4 mm between batches)	Detected at both levels; two-plateau structure	Conditional (session effect not excluded)
All gears beyond AGMA grade C12 (0.190 mm)	Best batch 0.220 mm; best gear 0.21 mm	Established (fixture contribution would make it conservative)
Mean ΔM of −0.37 mm (undersized gears)	Depends directly on uncertified ball diameter	Not established

These results can be compared with the gear studies examined, keeping in mind that the measured quantities differ. Egyed-Faluvégi et al. [19] applied a Box–Behnken design to PLA gears and found speed to be the only significant main effect on total profile deviation, with no clear layer effect. The two studies agree on a speed effect and differ on layer thickness, which is not a contradiction: profile deviation characterises the local form of the flanks, whereas runout reflects the position of the toothing, and their layer thickness was capped at 0.2 mm while the effect observed here continues to 0.3 mm. The dominance of layer thickness agrees with Karadag and Ulkir [21] for dimensional accuracy of printed gears and with specimen studies [11, 12].

4.2. Status of the dimension-over-balls results

A smaller dimension over balls indicates thinner teeth or a smaller toothing. Incomplete filling of the flanks at low temperature, where the melt is more viscous, or a more pronounced stair-step profile with thick layers would go in this direction, consistent with the two groups observed. A metrological explanation is, however, equally compatible with the data: the gears were measured on several days, and an offset between sessions (ball seating, micrometer

handling) would produce plateaus of the same kind. Since the allocation of gears to sessions was not recorded, the two explanations cannot be separated, and the effects of temperature and layer thickness on ΔM are reported as conditional. Re-measuring gears of both groups in a single session would settle the question.

The mean deviation of −0.366 mm means that all measured dimensions are smaller than M_0. Under the homothetic assumption, Eq. (6) would translate this into an equivalent pitch diameter about 0.44 mm below nominal. This value is not interpreted as an established undersizing: by Eq. (7), balls of 4.34 mm instead of 4.45 mm would cancel the mean deviation, and the ball diameter was not certified. More generally, the example shows why a single dimension over balls should be reported as such, with the assumption used for any conversion stated explicitly.

4.3. Position with respect to gear accuracy grades

Annex B of ANSI/AGMA 2015-2-A06 gives an informative runout tolerance for accuracy grades C4 (finest) to C12 (coarsest), applicable by agreement for $0.5 \le m_n \le 50$ and 5 mm $\le d \le$ 12 000 mm [28]. In micrometres,

$$F_{rT} = 0.8\,[0.025\,d + 0.3\,m_n + 19]\,\left(\sqrt{2}\right)^{C-5} \tag{13}$$

where C is the grade number. For the present gear the bracket equals 21.0, giving tolerances of 48, 67, 95, 134 and 190 µm for grades C8 to C12. None of the 45 gears reaches grade C12: the best gear (2a) has a runout of 0.21 mm and the best batch mean (batch 2: 0.1 mm, 220 °C, 120 mm/s) of 0.220 mm, 16% above the C12 limit, while 0.3 mm batches exceed it by a factor of 2.5–3.2 (Fig. 2). Because the standard defines grades C4–C12 only and excludes extrapolation, these gears are described as beyond grade C12 rather than assigned a fictitious grade. The reduced model predicts 0.207 mm (95% CI 0.187–0.228 mm) at the vertex (0.1 mm, 220 °C, 100 mm/s), i.e. at the C12 boundary, which gives the recommended confirmation runs an explicit acceptance criterion, $F_r \le 0.190$ mm. As the fixture may contribute to the measured runout, this classification is conservative. The 0.01 mm resolution of the dial indicator is about 5% of the C12 tolerance, adequate for this purpose as required by the standard, but would be insufficient for grades finer than about C9–C10. Finally, the standard notes that radial and elemental accuracy grades are not correlated, so this result says nothing about pitch or profile deviations.

4.4. Practical implications

Within the studied domain, layer thickness is the lever that matters for runout. Build times estimated by the slicer were about 38, 22 and 19 min per gear at 0.1, 0.2 and 0.3 mm layers. Moving from 0.2 to 0.3 mm saves about 3 min per gear but raises the marginal mean runout by about 0.13 mm (from 0.413 to 0.547 mm), which is rarely worthwhile for a gear; moving from 0.2 to 0.1 mm lowers it by about 0.16 mm (to 0.252 mm) for about 16 min more per gear. Raising the temperature towards 220 °C and using the lowest speed setting bring smaller additional gains. The batch-wise validation shows that a simple first-order model fitted to such a design predicts the mean runout of an untested configuration within a few hundredths of a millimetre.

4.5. Limitations

The study covers one material, one printer and one gear geometry, and the models are valid only inside the experimental domain. Only the centre configuration was replicated in separate builds, and gear positions on the plate were not recorded. No gauge repeatability and reproducibility study was performed; the shaft runout, the effect of mounting on the printed bore and the angular reference of tooth space 1 were not documented, so the fixture contribution to F_r is not quantified. For the dimension over balls, two readings instead of the three planned were taken, on several days with an unrecorded allocation, with uncertified balls. Seven overwritten readings on the measurement sheets remain to be confirmed. The slicer speed parameter was not identified. Other gear deviations (pitch, profile, helix) were not measured, and results were not related to tolerance classes. These limitations define the follow-up experiments: measuring the shaft runout alone, re-measuring ΔM in one session with calibrated balls, and printing confirmation gears at the predicted best setting.

5. Conclusions

A Box–Behnken experiment on 45 PLA spur gears printed in 15 batches, analysed at gear and batch levels, leads to the following conclusions.

- Radial runout is strongly process-driven: configurations explain 97% of its variance, against a within-batch SD of 0.019 mm.
- No gear reached the coarsest runout tolerance of ANSI/AGMA 2015-2-A06 (grade C12, 0.190 mm); only the thinnest layers approach it, and the predicted optimum lies at its boundary.

- Layer thickness explains about 85% of runout variation; runout increases with the speed setting and decreases moderately with nozzle temperature. A layer × speed interaction is detected but rests on four batches.
- The runout profile is dominated by a once-per-revolution component (median 99.6% of variance): the measured runout is an eccentricity of the toothing, whose origin (gear or fixture) remains to be determined.
- A first-order model fixed a priori predicts the mean runout of an unseen configuration with an RMSE of 0.025 mm in leave-one-batch-out validation; the PRESS statistic, which leaves sibling gears in the fit, is not a substitute.
- The deviation of dimension over balls varies strongly between configurations (96% of variance, repeatability 0.010 mm) but in two plateaus that a quadratic surface does not describe; its parameter effects are conditional on the absence of a measurement-session effect, and its absolute value depends on the ball diameter.

Treating the build rather than the part as the experimental unit, and reporting each conclusion with its strength of evidence, are simple practices that would make process studies of printed gears easier to compare and to reuse.

Declarations

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests: The authors declare no competing interests.

Ethics approval: Not applicable.

Consent to participate: Not applicable.

Consent for publication: All authors have read and approved the final manuscript and consent to its publication.

Data availability: The raw dial readings and the dimension-over-balls readings of the 45 gears are provided as electronic supplementary material.

Code availability: The Python scripts reproducing all tables and figures are provided as electronic supplementary material.

Author contributions: Fatiha Rabbaje: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft. Abdelmajid Ait Taleb: Supervision,

Methodology, Writing – review & editing. Larbi Lasri: Supervision, Validation, Writing – review & editing.

References

1. Gibson I, Rosen D, Stucker B (2015) Additive Manufacturing Technologies, 2nd ed. Springer, New York
2. Ngo TD, Kashani A, Imbalzano G, et al (2018) Additive manufacturing (3D printing): a review of materials, methods, applications and challenges. Compos B Eng 143:172–196. https://doi.org/10.1016/j.compositesb.2018.02.012
3. Bikas H, Stavropoulos P, Chryssolouris G (2016) Additive manufacturing methods and modelling approaches: a critical review. Int J Adv Manuf Technol 83:389–405. https://doi.org/10.1007/s00170-015-7576-2
4. Mašović R, Jagarčec V, Miler D, et al (2019) Analysis of printing direction impact on dimensional accuracy of spur gears. In: Mechanisms and Machine Science. Springer, pp 1111–1120
5. Buj-Corral I, Zayas-Figueras EE (2023) Comparative study about dimensional accuracy and form errors of FFF printed spur gears using PLA and Nylon. Polym Test 117:107862. https://doi.org/10.1016/j.polymertesting.2022.107862
6. Wei X (2025) Geometric accuracy and dimensional precision in 3D printing-based gear manufacturing: a study on interchangeability and forming precision. Polymers 17:416. https://doi.org/10.3390/polym17030416
7. Mohamed OA, Masood SH, Bhowmik JL (2015) Optimization of fused deposition modeling process parameters: a review of current research and future prospects. Adv Manuf 3:42–53. https://doi.org/10.1007/s40436-014-0097-7
8. Dey A, Yodo N (2019) A systematic survey of FDM process parameter optimization and their influence on part characteristics. J Manuf Mater Process 3:64. https://doi.org/10.3390/jmmp3030064
9. Sood AK, Ohdar RK, Mahapatra SS (2009) Improving dimensional accuracy of fused deposition modelling processed part using grey Taguchi method. Mater Des 30:4243–4252. https://doi.org/10.1016/j.matdes.2009.04.009

10. Mohamed OA, Masood SH, Bhowmik JL (2016) Mathematical modeling and FDM process parameters optimization using response surface methodology based on Q-optimal design. Appl Math Model 40:10052–10073. https://doi.org/10.1016/j.apm.2016.06.055

11. Stojković JR, Turudija R, Vitković N, et al (2023) An experimental study on the impact of layer height and annealing parameters on the tensile strength and dimensional accuracy of FDM 3D printed parts. Materials 16:4574. https://doi.org/10.3390/ma16134574

12. Frunzaverde D, Cojocaru V, Bacescu N, et al (2023) The influence of the layer height and the filament color on the dimensional accuracy and the tensile strength of FDM-printed PLA specimens. Polymers 15:2377. https://doi.org/10.3390/polym15102377

13. Frunzaverde D, Cojocaru V, Ciubotariu C-R, et al (2022) The influence of the printing temperature and the filament color on the dimensional accuracy, tensile strength, and friction performance of FFF-printed PLA specimens. Polymers 14:1978. https://doi.org/10.3390/polym14101978

14. Ardeljan DD, Frunzaverde D, Cojocaru V, et al (2025) The impact of elevated printing speeds and filament color on the dimensional precision and tensile properties of FDM-printed PLA specimens. Polymers 17:2090. https://doi.org/10.3390/polym17152090

15. Vanaei H (2020) Influence of process parameters on thermal and mechanical properties of polylactic acid fabricated by fused filament fabrication. Polym Eng Sci 60:1822–1831. https://doi.org/10.1002/pen.25419

16. Baechle-Clayton M, Loos E, Taheri M, Taheri H (2022) Failures and flaws in fused deposition modeling (FDM) additively manufactured polymers and composites. J Compos Sci 6:202. https://doi.org/10.3390/jcs6070202

17. Czégé L, Ruzicska G (2025) Geometrical analysis of 3D-printed polymer spur gears. Machines 13:422. https://doi.org/10.3390/machines13050422

18. Liu J, Feng K, Su Z, et al (2025) Dimensional deviations and distortion mechanism of polymer spur gear fabricated by fused deposition modeling. J Mater Eng Perform 34:3026–3035. https://doi.org/10.1007/s11665-024-09321-w

19. Egyed-Faluvégi E, Gergely AL, Kántor J (2025) Dimensional precision of plastic gear tooth profiles produced by 3D printing. Eng Proc 113:38. https://doi.org/10.3390/engproc2025113038

20. Malik ZH, Rifai AP, Muflikhun MA (2025) Comparative study of process parameters' influence on geometric accuracy and relative errors of lightweight compound gear printed by vat photopolymerization and material extrusion. J Eng Res 13:949–961. https://doi.org/10.1016/j.jer.2024.03.018

21. Karadag A, Ulkir O (2026) Experimental evaluation and ANN-based prediction of polymer spur gears fabricated by additive manufacturing. J Appl Polym Sci 143:. https://doi.org/10.1002/app.70535

22. (1997) ISO 1328-2:1997 — Cylindrical gears — ISO system of accuracy — Part 2: Definitions and allowable values of deviations relevant to radial composite deviations and runout information

23. Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. Ecol Monogr 54:187–211. https://doi.org/10.2307/1942661

24. Montgomery DC (2020) Design and Analysis of Experiments, 10th ed. Wiley, Hoboken, NJ

25. Box GEP, Behnken DW (1960) Some new three level designs for the study of quantitative variables. Technometrics 2:455–475. https://doi.org/10.1080/00401706.1960.10489912

26. Ferreira SLC (2007) Box-Behnken design: an alternative for the optimization of analytical methods. Anal Chim Acta 597:179–186. https://doi.org/10.1016/j.aca.2007.07.011

27. Bezerra MA, Santelli RE, Oliveira EP, et al (2008) Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta 76:965–977. https://doi.org/10.1016/j.talanta.2008.05.019

28. (2006) ANSI/AGMA 2015-2-A06 — Accuracy Classification System — Radial Measurements for Cylindrical Gears

29. Roberts DR (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40:913–929. https://doi.org/10.1111/ecog.02881

Supplementary material

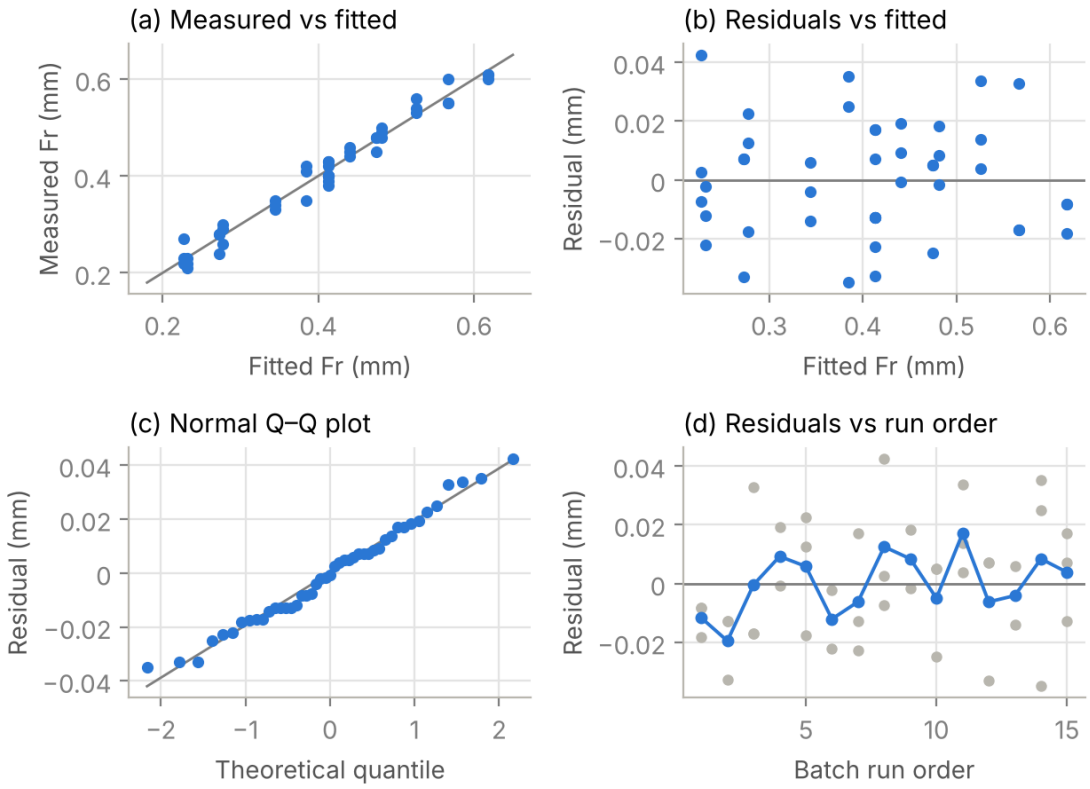


Fig. S1. Diagnostics of the reduced runout model (45 gears): (a) measured vs fitted; (b) residuals vs fitted; (c) normal Q–Q plot; (d) residuals vs batch run order.

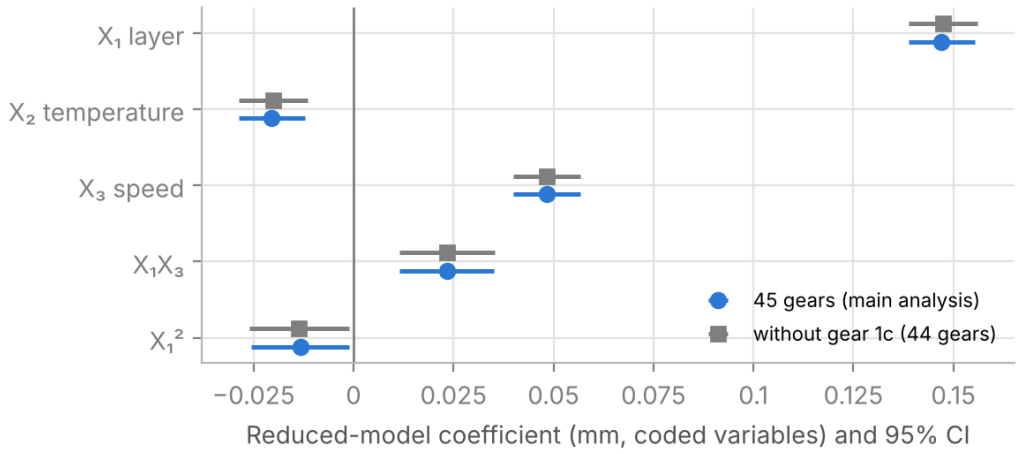


Fig. S2. Reduced-model coefficients and 95% confidence intervals with 45 gears and without gear 1c.

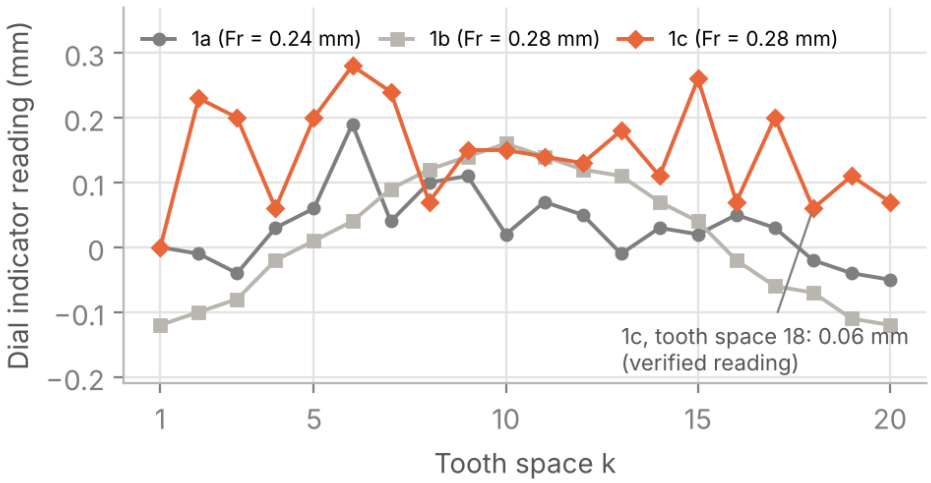


Fig. S3. Dial-indicator profiles of the three gears of batch 1 (0.1 mm; 190 °C; 120 mm/s); the verified reading of gear 1c in tooth space 18 is 0.06 mm.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExperimentalData45GearsRunoutM1M2.xlsx