recycling

MDPI

Article

Replacement Thresholds for Recycled Coarse Aggregate in Concrete: A Multi-Study Database from Peru

Sheila Schüler Pizan [1], José Ignacio Bendezú [1] and Emilio Jose Medrano-Sanchez [2,*]

[1] Civil Engineering, Universidad Tecnológica del Perú, Lima 15842, Peru; u18302339@utp.edu.pe (S.S.P.); u17214882@utp.edu.pe (J.I.B.)
[2] Faculty of Engineering, Universidad Tecnológica del Perú, Lima 15842, Peru
* Correspondence: c16602@utp.edu.pe

Abstract

Evidence on concrete with recycled coarse aggregate is abundant but scattered across unindexed documents that reach opposing conclusions for the same design strength, leaving designers without a Peruvian reference. This study quantified the association between the replacement ratio and mechanical performance of concrete with f′c = 210 kg/cm^2 (20.6 MPa) and established replacement thresholds, including their uncertainty. A database was assembled from Peruvian open-access repositories over a fifteen-year window, requiring a common design strength, a control mix, and complete test values; the nine primary studies that replaced coarse aggregate provided 179 compression and flexure observations at 7, 14, and 28 days. Each value was expressed relative to its own study's control and analyzed with Spearman correlation and random-intercept mixed-effects models. At 28 days, the association was negative and significant for compressive ($\rho = -0.695$; $\beta = -0.235$; $p < 0.001$) and flexural strength ($\rho = -0.760$; $\beta = -0.256$; $p < 0.001$); 23% to 51% of the variability lay between studies. Under an allowable loss of 10 percentage points, the prudent maximum replacement is 34% in compression and 29% in flexure; at 30% replacement, the estimated mean losses are 7.05 and 7.68 percentage points, and only the prudent calculation places flexure above that margin (10.26 points, against 8.94 in compression).

Keywords: recycled coarse aggregate; replacement ratio; relative strength; compressive strength; flexural strength; construction and demolition waste; circular economy

1. Introduction

Concrete is, after water, the most widely used material in the world, and aggregates account for up to 80% of its volume [1]. That composition explains why growth in construction translates directly into pressure on quarries: world concrete production quadrupled between 1990 and 2020, reaching 26 gigatonnes (Gt) per year, and meeting this production required roughly 20 Gt of virgin aggregate in 2020, an amount that exceeds the combined extraction of all fossil fuels on the planet [2]. As this extraction accelerates, the material capable of replacing it accumulates with no destination, because the concrete that reaches the end of its service life, some 4 Gt per year, is buried in landfills, remains locked in place, or is downcycled into road base, so that recycled aggregate accounts for less than 1% of all the aggregate the concrete industry consumes [2]. More stone than ever is therefore taken from quarries and riverbeds while the stone already extracted once is discarded.

This contradiction persists where the material has been studied and its use accepted: several national standards already set acceptable replacement percentages for structural

Check for updates

Academic Editor: Michele John

Received: 12 September 2026
Revised: 29 October 2026
Accepted: 3 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

concrete, ranging from 20% in Brazil and the United Kingdom to 30% in Australia and up to 35% in Germany for the lower strength classes [3]. The distance between what is permitted and what is practiced is nevertheless considerable. In the European Union, where the target of recovering 70% of construction and demolition waste was met in most countries, that compliance rested largely on backfilling operations that constitute a downgrade of the material and that, under certain conditions, amount to burying it, while recycled aggregate barely reaches 10.6% of all the aggregate produced there [1].

The authors of that review attribute the limited uptake to three circumstances: natural aggregate can be less expensive; clients, designers, and contractors still regard recycled material as unreliable; and, in many regions, processing plants do not produce aggregate suitable for concrete because no market exists for it [1]. All three converge at the point where the decision is made. Whoever signs off on a design is answerable for its performance, so distrust yields only to a figure showing how much can be replaced without compromising strength using what is at hand, which are the quarry, the cement, and the construction practices of each place. As long as that figure is missing for a given context, doing without the material remains the prudent option, even when its environmental merit is beyond dispute.

That individual prudence, repeated project after project, carries a collective cost: construction and demolition waste accounts for at least 30% of the solid waste generated worldwide, and its management is one of the central fronts of the circular economy in construction [4], so that the waste stream capable of returning to concrete is by a wide margin the largest of those that could reenter production. Its fate, however, depends less on its volume than on the capacity of each country to manage it; it is at this point that the problem ceases to be uniform.

In Latin America, the capacity is uneven and the waste frequently ends up where it should not. Brazil offers a measure of the mismatch: of the 79 million tonnes of municipal solid waste generated in 2018, 40.5% was dumped in inappropriate places, while the collection of construction waste was concentrated in the southeast region, which accounted for 52.1% of the total collected, against 3.9% for the north region [5]. Chile, which does quantify its generation, attributed 5,821,000 tonnes of waste to the construction sector in 2009 and has measured rates of 0.186 cubic meters per square meter built [6]. And the diagnosis of a Colombian municipality shows what happens when formal management does not reach the territory, since 58% of the waste is informally reused, 25% is abandoned in public spaces, and none of the sites visited held a building permit [7].

Peru shares those conditions and adds a shortfall of its own. Informal disposal is there the rule rather than the exception, since, of the 7.5 million tonnes of municipal solid waste generated in 2014, 47% ended up in open dumps and only 21% reached controlled landfills [8]. Regarding construction waste in particular, the national literature attributes the worsening of the problem to accumulation in illegal dumps, to deficient urban planning, and to limited citizen participation [9].

This waste is moreover produced on a building stock erected largely outside formal procedure, since self-construction has spread across Latin America as common practice, without prior studies, without permits, without professional supervision, and with materials whose quality goes unchecked [10]. Unlike the Brazilian and Chilean cases, no official statistic on construction and demolition waste generation is available for Peru, so that the magnitude of the national problem is assessed through its effects on the territory rather than through its measurement.

Against a problem of that magnitude, the technical response is far from incipient. Recycled coarse aggregate (RCA), obtained by crushing demolition concrete, left the question of its viability behind long ago, and international research has documented its behavior in

enough detail to explain why it performs below natural aggregate: the explanation lies in the layer of adhered mortar that coats the recycled particle, in a proportion that usually ranges between 20% and 60%, which gives rise to a more porous and weaker interfacial transition zone [11]. That understanding has been translated into magnitudes, since at full replacement compressive strength falls by about 12% and the modulus of elasticity by close to 18%, while flexural strength remains near that of the reference concrete [11].

With the mechanism established, the question in the field shifted from whether the material works to how much of it can be used, and at that point the answers ceased to be single. National standards set replacement limits ranging from 20% to 35%, depending on the country and the strength class [3], and experimental evidence shows that this limit depends on the strength of the concrete being produced, since the same replacement ratio behaves significantly differently in grade C30 concrete and in grade C50 concrete [12]. In moderate-strength concretes, close to 25 MPa (255 kg/cm^2), 10% replacement proves optimal and 20% compromises the homogeneity of the mix as well as its strength [13], while a recent review proposes 20% as the balance value for the body of available evidence [3]. Each of those limits answers to a different strength class, a different aggregate quality, and a different acceptance criterion, so that their dispersion measures the diversity of those contexts rather than an experimental disagreement.

What the Latin American region adds to the picture is not a technical discrepancy but a boundary condition. Construction and demolition waste is a heterogeneous material of undefined quality, so its conversion into aggregate requires processes that guarantee properties suitable for concrete [1]. To that variability is added the absence of specific rules and of uniform systems for identification, accreditation, and quality certification [14], and it is that absence which turns variability into an impediment because, without acceptance criteria of its own, each project has no way of deciding whether the material will serve its purposes. A review of North American frameworks has shown, in fact, that the void in policies and guidance aggravates technical reservations to the point of eliminating the option in many situations [15]. The obstacle, in other words, lies upstream of the test: what are missing are not measurements but the criteria with which to read them.

Downstream, by contrast, measurement has not stopped accumulating, and in Peru it has done so with a particular feature: the output on the subject is abundant, is recent, and converges moreover on the design strength most widely used in the country. Reference [16] obtained, for a concrete of f′c = 210 kg/cm^2 (20.6 MPa), a strength of 250.90 kg/cm^2 in the control concrete and of 220.44, 194.43, and 187.94 kg/cm^2 at replacements of 50%, 75%, and 100%. Reference [17], at the same design strength, found that 20% replacement exceeded the control concrete by 11.43% and 50% replacement by 4.45%, and that full replacement lost only 2.73%. Reference [18], also at f′c = 210 kg/cm^2, reported an average improvement of 16.63% at replacements of 20%, 25%, and 30%. None of the three is an isolated attempt, since national interest in replacing conventional constituents of concrete extends beyond recycled aggregate to other materials of residual origin, likewise assessed for their effect on compressive, tensile, and flexural strength [19]. What exists in the country is therefore a sustained line of work.

Read together, those studies describe incompatible behaviors for one and the same material at one and the same design strength, since the first finds growing losses beyond 50% replacement, the second finds gains up to that same percentage, and the third finds them across the entire range tested. Each is internally consistent and none departs from the test standards applied, so that the disagreement does not stem from a flaw in execution but from the variability inherent in working with different quarries, cements, waste origins, and curing conditions. A disagreement of that nature is not resolved by adding one more measurement to the set, which is what the field has been doing, but by integrating those

that already exist under a framework capable of separating how much of that variation corresponds to differences between studies and how much to the replacement ratio.

Integrating that evidence has already been attempted, and on a considerable scale. Reference [20] assembled 119 publications issued between 1978 and 2014 and subjected them to a joint analysis that made it possible to build a model for predicting concrete strength from the quality and the content of the recycled aggregate, with 95% confidence limits, while in related materials the random-effects meta-analysis applied to concrete with incineration bottom ash went as far as quantifying between-study heterogeneity [21]. Both integrations were built, however, on the published literature, and the second expressly excluded conference proceedings and book chapters; on that same class of record rests the most recent bibliometric map of the field, built from 1624 documents indexed in Scopus between 1983 and 2025, which located the dominant output in China, India, and the United States [22]. Since the image the field holds of itself likewise derives from what is indexed, the evidence that does not reach a journal remains outside both the syntheses and the map that guides them.

Against that body of evidence, several questions remain open. Replacement limits differ across standards and across reviews without always specifying to which strength class each one corresponds. Behavior in flexure rests on fewer studies than behavior in compression and is reported with less consistency. The quality of recycled aggregate still lacks uniform accreditation and certification criteria. And evidence that does not reach an indexed journal remains outside the syntheses, so that entire contexts of production go unintegrated. Of those questions, the present study addresses the last: it assembles the available Peruvian evidence under a common statistical framework, estimates the association between the replacement ratio and strength while separating the variation that corresponds to differences between studies, and translates that estimate into a threshold bounded by its uncertainty and referring to a specific strength class.

The Peruvian setting presents this question in its sharpest form. Primary evidence exists and is abundant; it concentrates moreover on a single design strength, $f'c = 210\ kg/cm^2$, which is the one most widely used in building construction in the country. Even so, it is scattered across professional-degree theses (the dissertation required for professional licensure in Peru) and conference papers that do not reference one another, to the point of supporting opposing conclusions about one and the same material.

To that dispersion is added a regulatory void that extends beyond the Peruvian case. Peruvian regulation sets the properties that recycled aggregate must meet and remains silent on the proportion in which it may be used [17], a silence it shares with the family of codes from which national practice derives, since the United States building code does not explicitly include recycled concrete [3]. An earlier review of those frameworks had shown that the silence runs wider, since neither did the structural specifications incorporate provisions on the use of recycled aggregate, nor did the aggregate standard set evaluation criteria of its own for the crushed concrete it did admit, so that the acceptance criteria coincided with those for natural aggregate [15]. The recent regional literature documents equivalent situations, whether because regulation prescribes preventive measures on the material rather than proportions of use [23], or because the scarcity of structural-grade data hinders the development of specifications of its own [24]. The percentages that other regulatory traditions do set fail to cover that void, since they answer to aggregates, cements, and practices different from Peruvian ones. As long as that silence persists, the replacement ratio will remain in the country a decision by the designer with no regulatory reference of its own to support it.

The Peruvian case is therefore of interest not as a geographic particularity but because in it the general absence takes a concrete and solvable form.

That this evidence resides in professional-degree theses, far from constraining the synthesis, enables it under favorable conditions. Searching the gray literature systematically is an established procedure, adaptable to different disciplines, even though no reference standard for carrying it out yet exists [25], and the source offers two advantages of its own here. Theses are deposited independently of the result obtained, so that a set built from them is less exposed to publication bias than one made up only of articles. And they preserve the laboratory documentation for which the length limits of a journal leave no room, so that every test value remains verifiable at its source.

Bringing those sources together into a single database requires first overcoming the obstacle that prevents comparing them with one another, and it is there that the design does its work. Because each study starts from a control strength of its own, conditioned by its quarry, its cement, and its curing conditions, absolute values admit no direct comparison; expressing each measurement as a proportion of the control concrete of its own study, that is, as relative strength, returns them to a common scale and neutralizes that initial heterogeneity. On that scale, a mixed-effects model incorporating the study of origin as a random intercept makes it possible to distinguish how much variation comes from differences between studies and how much from the replacement ratio, and on that distinction depends whether the result can be expressed as a threshold rather than as an average. Since normalization removes the differences in absolute strength between sources, the same synthesis can be repeated in any context where the evidence is equally scattered, and in each such context it yields a threshold of its own.

From that framing follows the question that guides this study: to what extent the replacement ratio of natural coarse aggregate by RCA is associated with the mechanical performance of concrete of $f'c = 210\ kg/cm^2$, and what maximum replacement follows from that association once its uncertainty is taken into account. Translated into operational terms, the objective is to determine and quantify the extent to which this replacement ratio is associated with the mechanical performance of the concrete, in order to establish replacement thresholds with their associated uncertainties.

That objective unfolds along the two dimensions in which mechanical performance is examined, compressive strength and flexural strength, assessed at 7, 14, and 28 days, and from that structure the hypotheses of the study are derived. The general hypothesis holds that the replacement ratio of natural coarse aggregate by RCA is significantly associated with the mechanical performance of the concrete. The first specific hypothesis holds that this ratio is significantly associated with compressive strength, and the second, that it is significantly associated with flexural strength.

What is decided in answering these hypotheses exceeds the technical plane and is organized around Sustainable Development Goal 12, responsible consumption and production, whose purpose is to decouple economic growth from the consumption of natural resources. Construction concentrates that decoupling at a very precise point, since every tonne of recycled aggregate incorporated into a mix is a tonne no longer taken from a quarry and a tonne that no longer accumulates in a dump. That exchange depends, however, on the existence of a percentage backed by local evidence on which to base a design, so that a threshold accompanied by its uncertainty transcends the technical plane: it is the condition that allows the decision to recycle to stop resting on the individual disposition of whoever signs and to rest instead on a verifiable criterion.

2. Results

This chapter sets out the results of the statistical analysis applied to the 179 relative strength observations extracted from the nine primary studies. Each observation is expressed as relative strength, that is, as a percentage of the control concrete of its own study,

which therefore equals 100%; the database, the normalization and the models are described in Section 4. The findings are organized in a logical sequence that advances from the descriptive to the inferential, in which each section prepares the next: the behavior of relative strength is first described by the type of test and age; from that description, the normality assumption is verified, which determines the relevant correlation test; with the test thus justified, the association between the replacement ratio and strength is examined; that association is then quantified by a linear mixed-effects model that controls the dependence among observations from a single study; and finally the hypotheses are tested in the light of the evidence obtained. The complete statistical output supporting this chapter is presented in the supplementary material (Supplementary S5).

The analysis begins with the composition of each subgroup and its descriptive characterization, a step needed to establish the magnitude and the variability of the phenomenon before putting it to the test.

2.1. Coverage and Descriptive Analysis of Relative Strength

Table 1 shows how many studies and observations make up each combination of type of test and test age; each age gathers its own set of studies and replacement ratios, so each subgroup is read on its own. Table 2 presents the descriptive statistics of relative strength for those subgroups. In compression, the median exceeded the mean at all three ages, which shows that most mixes kept a strength close to that of their control, while between a quarter and a third of the observations, mostly at replacement ratios of 40% or more, lost more than ten points and pulled the mean down. At 28 days, that subgroup also had the lowest standard deviation of the compression subgroups, so it is where the studies lie closest to one another. In flexure, three of the fifteen recycled mixes tested at 7 and at 14 days matched or exceeded their control, and those two subgroups had the lowest standard deviations of the six, 6.21 and 5.95. At 28 days, the flexural median was below the control and above the mean, so most mixes lost strength and four values below 80% pulled the average down. In all subgroups, the number of missing datapoints was zero, and the number of observations per subgroup ranged between 19 and 40.

Table 1. Coverage of the included sources by type of test and test age.

Test	Age (Days)	*k*	*n*
Compression	7	8	35
Compression	14	7	31
Compression	28	9	40
Flexure	7	4	19
Flexure	14	4	19
Flexure	28	8	35

Note: *k* = number of independent studies contributing observations to the subgroup; *n* = number of observations extracted. Prepared from the cleaned analytical database (Supplementary S3).

Table 2. Descriptive statistics of relative strength (%) by type of test and test age.

Test and Age	*n*	M	Mdn	SD	Min.	Max.
Compression 7 d	35	93.5	97.4	9.52	67.1	106
Compression 14 d	31	92.5	94.8	10.07	70.8	109
Compression 28 d	40	93.1	96.1	8.51	71.5	103
Flexure 7 d	19	96.4	97.3	6.21	78.3	105
Flexure 14 d	19	96.5	97.3	5.95	80.7	108
Flexure 28 d	35	91.4	92.7	10.47	60.7	110

Note: *n* = number of observations; M = mean; Mdn = median; SD = standard deviation; Min. = minimum; Max. = maximum. Relative strength is expressed as a percentage of the control concrete (0% replacement) in each study.

These statistics describe the central tendency and the variability of each subgroup, but on their own they do not establish that the replacement ratio bears on strength: a mean below the control suggests a loss, yet it does not prove that the loss is statistically attributable to replacement. Establishing that association calls for inferential analysis. The choice of the appropriate correlation test, however, depends on whether the data satisfy the normality assumption, whose verification is taken up in the following section.

2.2. Verification of the Normality Assumption

In order to determine the relevant correlation test, the normality of relative strength was assessed in each subgroup by the Shapiro–Wilk test, chosen because its power suits the number of observations in each subgroup (between 19 and 40). The results, presented in Table 3, indicate that, in four subgroups, compression at the three ages and flexure at 28 days, the p values were below 0.05, ranging between less than 0.001 and 0.009, so that the normality hypothesis is rejected in them; in flexure at 7 and 14 days, the p values were 0.051 and 0.462. This condition was visually corroborated by the quantile–quantile plots, in which the values departed from the theoretical diagonal. The Spearman correlation coefficient is therefore applied to the six subgroups, so that all of them are compared with the same coefficient.

Table 3. Shapiro–Wilk normality test for relative strength by test and age.

No.	Test and Age	*n*	W	*p*
1	Compression 7 d	35	0.844	<0.001
2	Compression 14 d	31	0.893	0.005
3	Compression 28 d	40	0.870	<0.001
4	Flexure 7 d	19	0.901	0.051
5	Flexure 14 d	19	0.954	0.462
6	Flexure 28 d	35	0.913	0.009

Note. W = Shapiro–Wilk statistic; p = probability value. A p value < 0.05 indicates significant departure from normality.

Figure 1 shows the departure in the compression subgroups and in flexure at 28 days: the points separate from the theoretical diagonal in the tails and form a step in the central zone, where the control concrete values accumulate. In flexure at 7 and 14 days, the points follow the diagonal except for the lowest value.

2.3. Correlation Analysis Between the Replacement Ratio and Relative Strength

Once the relevance of the nonparametric test was confirmed, the Spearman correlation coefficient between the replacement ratio and relative strength was computed in each subgroup. Table 4 summarizes the coefficients obtained, their significance, and the corresponding number of observations.

Table 4. Spearman correlation coefficients between the replacement ratio and relative strength.

No.	Test and Age	*n*	ρ	*p*
1	Compression 7 d	35	−0.633	<0.001
2	Compression 14 d	31	−0.595	<0.001
3	Compression 28 d	40	−0.695	<0.001
4	Flexure 7 d	19	−0.681	0.001
5	Flexure 14 d	19	−0.714	<0.001
6	Flexure 28 d	35	−0.760	<0.001

Note: n = number of observations; ρ = Spearman correlation coefficient; p = probability value. A p value < 0.05 indicates a statistically significant correlation.

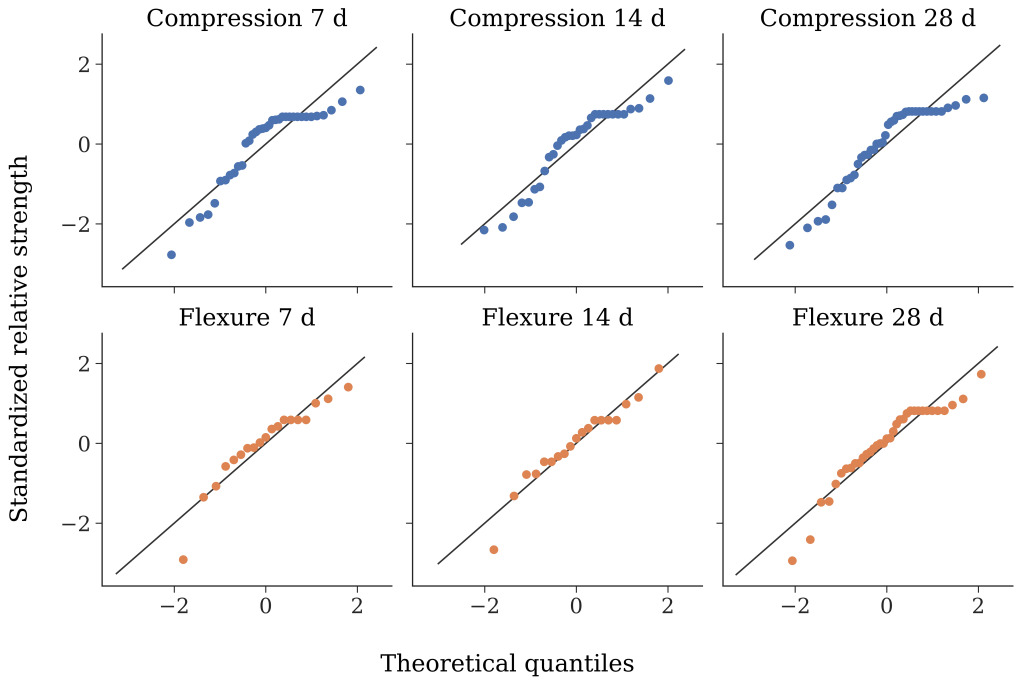


Figure 1. Quantile–quantile (Q–Q) plots of relative strength by type of test and age. Note: The diagonal line represents the theoretical normal distribution. The departure of the points from the diagonal, as well as the stepped stretches in the central zone (attributable to the concentration of values at 100%, corresponding to the control concretes), evidences the departure from normality.

In the compression test, the results show a negative and significant association at the three ages, moderate at 7 and 14 days ($\rho = -0.633$ and $\rho = -0.595$, $p < 0.001$) and strong at 28 days ($\rho = -0.695$, $p < 0.001$), the subgroup that gathers all nine studies.

In the flexure test, the negative association is strong and highly significant at all three ages ($\rho = -0.681$ at 7 days, $\rho = -0.714$ at 14 days, and $\rho = -0.760$ at 28 days, all with $p \leq 0.001$), including the two early ages, which rest on four studies each. That consistency agrees with the role played by the interface between the paste and the aggregate in that property. How much strength is lost per point replaced is, by contrast, a question of magnitude that correlation does not settle and that the mixed-effects model estimates further on.

Taken together, the six subgroups confirm a negative association between the replacement ratio and relative strength, that is, the greater the replacement of natural aggregate by recycled aggregate, the lower the strength retained relative to the control concrete, at every age and in both tests.

Figure 2 allows that association to be followed observation by observation: in the six subgroups, the cloud descends as replacement increases, with lines of similar slope in the three compression subgroups, flatter ones in flexure at 7 and 14 days, and the steepest one in flexure at 28 days.

These correlations, however, treat each observation as independent, without allowing for the fact that several come from a single study. Their magnitude is therefore confirmed and refined by the linear mixed-effects model, which incorporates explicitly the dependence among observations from a single source.

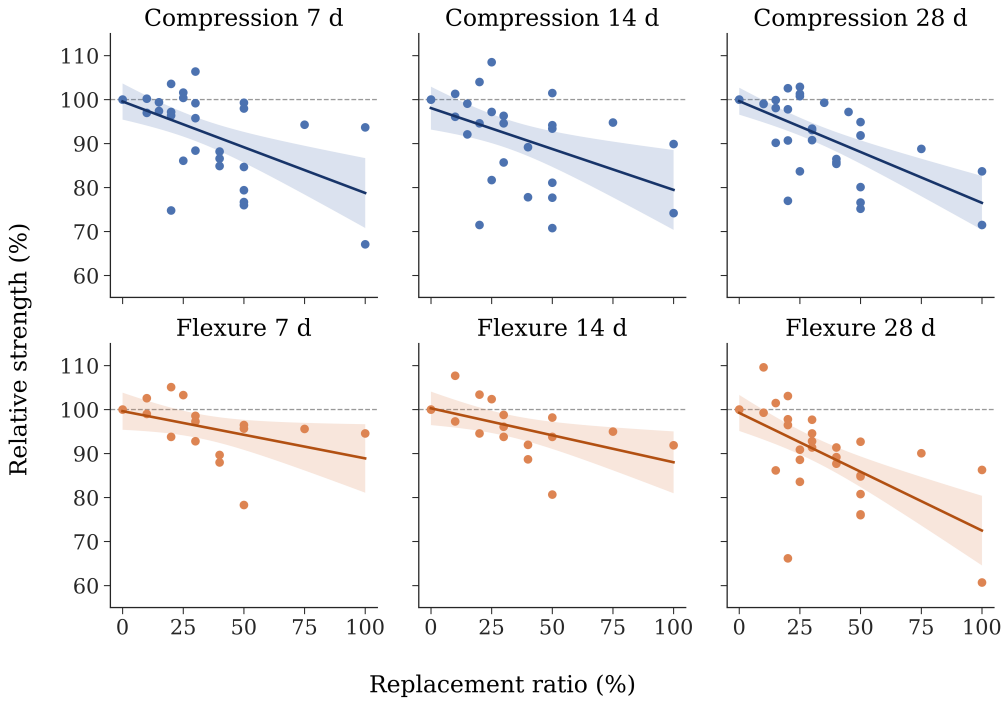


Figure 2. Scatterplots of relative strength against the replacement ratio, by type of test and age. Note: Each point represents an observation and each line the simple linear trend of the subgroup, with its 95% confidence band; these lines do not incorporate the dependence among observations from a single study, which the mixed-effects model in Section 2.4 does control. The descending slope of the lines confirms the negative association between the replacement ratio and relative strength in all six subgroups. The horizontal dashed line marks the level of the control concrete (100%).

2.4. Linear Mixed-Effects Model

To estimate the magnitude of the effect of the replacement ratio while controlling the dependence among observations from a single study, a linear mixed-effects model was fitted in each subgroup, with relative strength as the dependent variable, the replacement ratio as a fixed effect, and the study of origin as a random intercept. The six models reached convergence, and Table 5 presents the fixed-effect estimate, its 95% confidence interval, its significance, and the intraclass correlation coefficient (ICC) for each subgroup.

The intraclass correlation coefficients ranged between 0.225 and 0.510, which indicates that between 23% and 51% of the total variability in relative strength is attributable to differences between studies. This magnitude confirms that the dependence among observations from a single source was substantial, above all at 28 days (0.510 in compression and 0.445 in flexure), and supports the use of the mixed-effects model over an analysis assuming independence.

As for the fixed effect, the model estimated a negative slope in all six subgroups, consistent with the direction of the correlations. In compression, the effect was significant at the three ages, at 7 days ($\beta = -0.221$, 95% CI [−0.311, −0.131], $p < 0.001$), at 14 days ($\beta = -0.218$, 95% CI [−0.314, −0.122], $p < 0.001$), and at 28 days ($\beta = -0.235$, 95% CI [−0.298, −0.173], $p < 0.001$). The interpretation of the coefficient at 28 days indicates that, for each additional percentage point of replacement, relative compressive strength decreases by 0.235 percentage points; a 50% replacement is accordingly associated with an estimated loss of approximately 11.8 percentage points relative to the control concrete. In flexure, the effect was significant at the three ages as well, at 7 days ($\beta = -0.132$, 95% CI [−0.234, −0.030], $p = 0.014$), at 14 days ($\beta = -0.148$, 95% CI [−0.237, −0.059], $p = 0.003$), and at 28 days ($\beta = -0.256$, 95% CI [−0.342, −0.169], $p < 0.001$). Repeated at 28 days without the control observations, the slope remained negative and significant in compression ($\beta = -0.218$,

95% CI [−0.297, −0.140], $p < 0.001$) and in flexure ($\beta = -0.199$, 95% CI [−0.309, −0.088], $p = 0.001$), as did the Spearman correlation, at $\rho = -0.562$ and $\rho = -0.589$, with $p = 0.001$ in both cases (Supplementary S4.7 and S4.8).

Table 5. Linear mixed-effects model estimates for the effect of the replacement ratio, by test and age.

No.	Test and Age	*k*	β	95% CI	*p*	ICC	R^2_m	R^2_c	*p* (LRT)
1	Compression 7 d	8	−0.221	[−0.311, −0.131]	<0.001	0.390	0.341	0.598	0.014
2	Compression 14 d	7	−0.218	[−0.314, −0.122]	<0.001	0.492	0.301	0.645	0.004
3	Compression 28 d	9	−0.235	[−0.298, −0.173]	<0.001	0.510	0.459	0.735	<0.001
4	Flexure 7 d	4	−0.132	[−0.234, −0.030]	0.014	0.225	0.270	0.435	0.329
5	Flexure 14 d	4	−0.148	[−0.237, −0.059]	0.003	0.308	0.360	0.557	0.183
6	Flexure 28 d	8	−0.256	[−0.342, −0.169]	<0.001	0.445	0.406	0.670	0.007

Note: *k* = number of studies; β = fixed-effect estimate of the replacement ratio (change in relative strength per percentage point of replacement); 95% CI = confidence interval; ICC = intraclass correlation coefficient; R^2_m = marginal R^2, variance explained by the replacement ratio; R^2_c = conditional R^2, variance explained by the replacement ratio and the study together; *p* (LRT) = probability of the likelihood-ratio test comparing each model with the same model without the random intercept, that is, an ordinary least squares regression. All models reached convergence. The flexure subgroups at 7 and 14 days rest on only 4 studies, so their estimates are interpreted with caution.

The two analyses agree. In the six subgroups, the mixed-effects model confirms the significance found in the correlation, including flexure at 7 and 14 days, which rest on four studies each. Taken together, the negative association between replacement and mechanical performance holds at the three ages, in both compression and flexure, once the dependence between studies is controlled. The fit supports that control in the three compression subgroups and in flexure at 28 days, where the random intercept improved on an ordinary least squares regression with *p* between <0.001 and 0.014; in flexure at 7 and 14 days, that test gave $p = 0.329$ and 0.183, and in all six subgroups the conditional R^2, between 0.435 and 0.735, exceeded the marginal R^2, between 0.270 and 0.459 (Table 5). The random intercept alone was retained at 28 days, since allowing each study a slope of its own did not improve the fit in compression ($p = 0.143$, $\beta = -0.240$) or in flexure ($p = 0.284$, $\beta = -0.255$), and that model returned a singular fit in compression and did not converge in flexure. These tests show that the data do not support study-specific slopes, without establishing that all studies share the same slope. In flexure, the third-point sources alone gave a slope of −0.239, 95% CI [−0.346, −0.133], close to the −0.256 of the complete set (Supplementary S7); their comparison with the other two sources, whose documents report no loading method, gave a difference of 0.118 ($p = 0.371$).

2.5. Hypothesis Testing

The hypotheses are tested on the evidence obtained. The testing takes 28 days as its reference, since that is the normative age at which the design strength of concrete is specified, and it integrates the two inferential analyses applied: the Spearman correlation, which indicates the direction of the association, and the linear mixed-effects model, which confirms its significance once the dependence between studies is controlled. Table 6 summarizes the verdict for each hypothesis.

The first specific hypothesis holds that the replacement ratio of natural coarse aggregate by RCA is significantly associated with compressive strength. The results at 28 days support this hypothesis, since the correlation was negative and significant ($\rho = -0.695$, $p < 0.001$) and the mixed-effects model confirmed a significant effect ($\beta = -0.235$, $p < 0.001$). The first specific hypothesis is therefore accepted.

Table 6. Summary of the hypothesis testing.

Hyp.	Statement	Evidence (28 Days)	Decision
SH1	The replacement ratio is significantly associated with compressive strength.	$\rho = -0.695$ ($p < 0.001$) $\beta = -0.235$ ($p < 0.001$)	Accepted
SH2	The replacement ratio is significantly associated with flexural strength.	$\rho = -0.760$ ($p < 0.001$) $\beta = -0.256$ ($p < 0.001$)	Accepted
GH	The replacement ratio is significantly associated with the mechanical performance of the concrete.	Supported by SH1 and SH2	Accepted

Note: GH = general hypothesis; SH = specific hypothesis; ρ = Spearman correlation coefficient; β = fixed effect of the mixed-effects model. The statistics correspond to 28 days, the normative reference age. The decision considers the joint evidence of both analyses.

The second specific hypothesis holds that this same ratio is significantly associated with flexural strength. The results at 28 days support it as well, with a negative and significant correlation ($\rho = -0.760$, $p < 0.001$) and a significant effect in the mixed-effects model ($\beta = -0.256$, $p < 0.001$). The second specific hypothesis is therefore accepted.

Finally, since both specific hypotheses are accepted, and since compressive strength and flexural strength constitute the dimensions of the mechanical performance of concrete, the general hypothesis is accepted: the replacement ratio of natural coarse aggregate by RCA is significantly associated with the mechanical performance of the concrete. The association is negative in direction, that is, an increase in the replacement ratio is related to a decrease in strength relative to the control concrete.

In summary, the statistical evidence supports the three hypotheses at the reference age, confirming that the incorporation of RCA is associated with a significant reduction in the mechanical performance of the concrete. The two slopes at 28 days are of similar magnitude, 0.235 in compression and 0.256 in flexure, and what sets flexure apart is the precision of its estimate, since its confidence interval is about 40% wider than that of compression.

2.6. Estimation of Replacement Thresholds

With the hypotheses tested, the magnitude estimated by the mixed-effects model admits a reading of direct use. Since the fixed effect expresses the change in relative strength per percentage point of replacement, the expected loss for a given ratio is obtained by multiplying that ratio by the point estimate, and its uncertainty propagates in the same way from the limits of the confidence interval. Table 7 presents that estimate at 28 days, which is the normative reference age. The linear slope is retained to describe that loss across the range studied, since the quadratic term in the replacement ratio was not significant in either property at 28 days ($p = 0.243$ and $p = 0.084$; Supplementary S4.3 and S4.6). That result shows that the data do not support a curved form, without establishing that the relationship is strictly linear.

The reading of those values depends on the criterion adopted, and the inverse calculation makes it explicit: the maximum replacement compatible with a given allowable strength loss follows from dividing that loss by the absolute value of the corresponding coefficient. Table 8 applies that calculation to an allowable strength loss of 10 percentage points under the three scenarios that the confidence interval delimits, and the supplementary material extends the same derivation to other allowable losses (Supplementary S5).

On the point estimate, that margin is exhausted only at a replacement of 43% in compression and 39% in flexure. A design decision, however, is taken on the unfavorable scenario and not on the central value, and, under the confidence-interval endpoint corresponding to the greater strength loss, the same margin is preserved only up to 34% in compression and up to 29% in flexure.

Table 7. Estimated loss of relative strength by replacement ratio, at 28 days.

Replacement (%)	Compression: Loss (pp)	95% CI	Flexure: Loss (pp)	95% CI
10	2.35	[1.73, 2.98]	2.56	[1.69, 3.42]
20	4.70	[3.46, 5.96]	5.12	[3.38, 6.84]
30	7.05	[5.19, 8.94]	7.68	[5.07, 10.26]
40	9.40	[6.92, 11.92]	10.24	[6.76, 13.68]
50	11.75	[8.65, 14.90]	12.80	[8.45, 17.10]

Note: pp = percentage points of loss relative to the control concrete of each study. The estimates derive from the fixed effect of the linear mixed-effects model at 28 days ($\beta = -0.235$ in compression and $\beta = -0.256$ in flexure), and their interval propagates from the limits of the fixed effect. They are neither independent measurements nor test results.

Table 8. Maximum replacement compatible with a loss of ten percentage points, by scenario.

Scenario	Compression (%)	Flexure (%)
Point estimate of the coefficient	42.6	39.1
95% CI endpoint corresponding to the greater strength loss (prudent criterion)	33.6	29.2
95% CI endpoint corresponding to the smaller strength loss (favorable scenario)	57.8	59.2
95% CI endpoint corresponding to the greater strength loss, with the source that added recycled aggregate in proportion to the cement mass	33.8	25.6

Note: The values are obtained by dividing the allowable strength loss of 10 percentage points by the absolute value of the corresponding coefficient from Table 5; the last row repeats the prudent criterion with the source that added recycled aggregate in proportion to the cement mass (Supplementary S4.9 and S4.10). The estimates correspond to the mean slope of the set of studies and not to the variability of an individual mix.

In the source that added recycled aggregate in proportion to the cement mass, the mixes with 5%, 10%, and 15% added aggregate exceeded their control at every age and, with 15%, reached 106.5% in compression and 153.0% in flexure at 28 days. With that source, the prudent threshold is 34% in compression and 26% in flexure (Table 8).

Within the nine sources, weighting each observation by the number of specimens behind its mean leaves the direction and significance of the slope unchanged in the six subgroups and places those limits at 35% and 28% (Supplementary S6), with the same order between the two properties. Weighting instead by the standard deviations that the sources report, on the observations that report them, lowers the prudent threshold from 34.0% to 28.9% in compression and from 28.5% to 16.8% in flexure, with two studies carrying 91% and 92% of the total weight; setting either of them aside places the threshold between 22.5% and 66.5% in compression and between 14.6% and 22.2% in flexure (Supplementary S6). The prudent threshold remains at 34% in compression and 29% in flexure when the interval is computed with Kenward–Roger degrees of freedom or by parametric bootstrap (Supplementary S4.3 and S4.6). Returning to the analysis the records left out of the database—the flexure of one included source and the two sources set aside in full—raises it to 38% in both properties (Supplementary S4.11 and S4.12), so that the thresholds adopted correspond to the more prudent scenario. Setting aside the sources with less documented quality (Section 4.2)—those without attached certificates or without a declared test standard—places the prudent threshold between 34% and 38% in compression and between 30% and 34% in flexure (Supplementary S7). In all these analyses, the slope remains negative and significant. Flexure thus governs the threshold because its interval is wider, a width associated with the smaller number of studies and the greater dispersion of the subgroup, which, together with a slightly steeper slope, pushes its confidence-interval endpoint corresponding to the greater strength loss further out; the scarcer evidence thus translates into a more prudent threshold.

That dependence between the amount of evidence and the severity of the threshold also delimits the scope of the values obtained. Deriving from the mean slope of the set of studies, they describe the aggregate behavior of the material and not the dispersion proper for a particular mix, so they guide the choice of the replacement ratio without replacing the testing of the aggregate each project uses. With that scope defined, the estimated association is translated into a scale for decision, whose practical reading is taken up in the following chapter.

3. Discussion

3.1. Replacement Thresholds Against the Literature and the Code

The results admit a unified reading once they are ordered around the question that motivated the study, which is not whether RCA alters the mechanical performance of concrete, settled decades ago, but how far replacement can be taken without compromising the design strength. Each finding contributes one condition to that threshold: that a usable range exists, that it cannot be read from an isolated study, at what age and with which property it is set, and what its figure is.

The estimate of this study sets the loss of relative compressive strength at 0.235 percentage points per point of natural aggregate replaced, and the loss in flexure at 0.256 points, both at 28 days and with the confidence intervals recorded in Table 5. That the fall is proportional to replacement in both properties, rather than an abrupt drop beyond a given ratio, is what makes it possible to speak of a threshold and not of a failure limit, because performance at any intermediate level is anticipated by the slope itself.

The Peruvian literature responds to that result in two opposing ways. Reference [16] agrees, since its series falls from 250.90 kg/cm^2 in the control concrete to 187.94 kg/cm^2 at full replacement, which amounts to between 0.24 and 0.30 percentage points of loss per point replaced, a range that overlaps the interval estimated here. References [17,18] disagree, since they report increases over the control of 11.43% at 20% replacement and of 16.63% across the 20% to 30% range. The difference lies in the level of aggregation, since a mean slope estimated over nine studies is compatible with an individual measurement that exceeds its own control. What the association establishes is therefore that, between zero and full replacement, there exists a range in which the loss remains bounded and predictable, and that range is the object of everything that follows.

Delimiting that range requires first knowing where the variation that obscures it resides, and the second result locates it. The mixed-effects models fitted in this study attribute between 23% and 51% of the variability in relative strength to differences between studies, a proportion that survives normalization, since the values had already been expressed as a proportion of the control of each study and the scale discrepancy between control mixes was discounted beforehand. What remains after that discounting brings together the material, the practice of each laboratory and the way it measures, and its magnitude means that, at 28 days, roughly half of what is observed when two experiments are compared does not come from the replacement ratio.

The literature confirms that result along two routes. The first is mechanical, since every recycled particle carries adhered mortar in a proportion that usually ranges between 20% and 60%, which introduces a weak interfacial transition zone [11], so that two studies declaring the same ratio may have replaced different materials; the strength class, which also conditions the effect [12], is ruled out as an explanation because the design held it constant. The second is the disagreement among the sources, visible in that [16–18] share a design strength and reach incompatible conclusions, and in that [13] ruled out 20% at a nearby strength; this is that same finding that led [20] to build their model on 119 publications and [21] to quantify heterogeneity in a related material. The implication is

twofold, since the threshold cannot be read from one experiment and any limit derived from this analysis will have to be presented with a band and not as a single value.

The third result indicates at what age that threshold should be set, and it points to 28 days. In compression, the slope of the mixed-effects model is significant at all three ages, at −0.221, −0.218, and −0.235. In flexure it is significant at all three ages as well, at −0.132, −0.148, and −0.256, with four studies at 7 and at 14 days. Twenty-eight days is the age at which strength is submitted for acceptance and the one at which both properties gather the largest number of studies.

With the age settled, the next question is which of the two properties governs the limit. The slopes at 28 days are of similar magnitude, 0.235 in compression and 0.256 in flexure, and what sets them apart is precision, since the flexural interval, 95% CI [−0.342, −0.169], is about 40% wider than that of compression, [−0.298, −0.173], and that width, together with the slightly steeper slope, is what lowers the prudent threshold in flexure to 29%. That width has an identifiable origin in the smaller number of studies and in the greater dispersion of the subgroup, with a standard deviation of 10.47 against 8.51, while the third-point sources alone (Section 4.2) yield a slope close to that of the complete set.

The literature offers here the most demanding contrast, since [11] reports that recycled aggregate did not affect flexural strength while compression fell 12.2% at full replacement. That result is not contained within the interval estimated in this study, which excludes zero, although neither does it contradict it, because the interval describes the mean slope and not the behavior of an individual experiment: with 45% of the variability attributed to differences between studies in the flexure subgroup at 28 days, an isolated study may record no loss at all without the mean slope ceasing to be negative. The consequence for the threshold is that it is governed by the property estimated with the least precision, so that the scarcer evidence is paid for with the more prudent threshold.

The fourth result translates the slope into the figure the question demanded. Under an allowable strength loss of 10 percentage points relative to the control, the point estimate preserves that margin up to 43% replacement in compression and 39% in flexure, whereas, under the prudent criterion of the confidence-interval endpoint corresponding to the greater strength loss, the one that applies to a design decision, the threshold falls to 34% and 29%. The answer is those two numbers and not one, because the admissible replacement depends on which property governs the element, and flexure restricts it because its slope is estimated with less precision.

The allowable strength loss of 10 percentage points, on which the entire threshold depends, is a criterion adopted and declared by this study rather than a value established by regulation, and shifting it relocates the threshold proportionally, since at 5 points the prudent threshold would fall to 17% in compression and 15% in flexure, and at 15 points it would rise to 50% and 44%. Declaring it in this way turns the result into a decision function, so that anyone designing under a different criterion can recompute their own limit from the slope and the interval this study provides, which is precisely what no Peruvian source offered until now. The thresholds reported here therefore operate as statistical references conditional on that criterion and on the specification of the analysis, since they derive from the confidence interval of the mean slope across studies and not from a safety bound for an individual mixture. At every age and under every specification examined at 28 days, which vary the estimation of the interval, the sources retained, and the weighting of the observations, the slope remains negative and significant, whereas the prudent threshold ranges between 23% and 67% in compression and between 15% and 38% in flexure (Section 2.6). The principal analysis gives every observation the same weight, since most standard deviations come from three or fewer specimens, they are available

for 88% of the observations at 28 days, and weighting by them places 91% and 92% of the weight on two studies.

The allowable strength loss of 10 percentage points was read against the provision that Peruvian regulation makes for production variability. When the producer has no records from which to establish a standard deviation, Norma Técnica E.060 (the Peruvian reinforced concrete code) sets in its Table 5.3 the required average strength as $f'cr = f'c + 7.0$ MPa (71 kg/cm^2) for specified strengths below 21 MPa (214 kg/cm^2) [26], so that a mix of $f'c = 210$ kg/cm^2 must be proportioned to reach around 281 kg/cm^2. That margin exists so that the dispersion between batches keeps individual results above the specified strength, and this study does not treat it as an allowance available to absorb the reduction associated with recycled aggregate. It is reported here as a consistency check: applied to that proportioning, a loss of 10 percentage points leaves a mean of about 253 kg/cm^2, above the specified value, while the protection against batch-to-batch dispersion is reduced in the same measure. The controls of the nine studies reached between 224 and 385 kg/cm^2 at 28 days, with a mean of 285.7 kg/cm^2, so that the same 10 percentage points correspond in each study to an absolute loss of between 22.4 and 38.5 kg/cm^2. The provision governs compressive strength, and the code offers no equivalent basis for an allowable loss in flexure, so the criterion adopted for that property rests on the declared decision of this study alone.

Set against international regulation, the threshold obtained proves less bold than it appears. The limits in force form a band between 20% and 35%, within which Australia admits 30%, Brazil and the United Kingdom 20%, and Germany up to 35% in its lower strength classes [3]. The 34% derived here for compression falls within that band and very close to its upper end, a convergence that gains value from being independent, since it comes from different regulatory traditions and from experimental evidence this work did not use. The comparison is established with the band and not with a specific class, since European classes are defined on cube specimens and the strength specified here on cylinder specimens.

Within that framework, the conclusion in [3], which proposes 20% as the optimal replacement, derives from the whole body of global evidence and from all strength classes at once, whereas [12] showed that the effect of replacement depends on the class, which is why the comparison is legitimate only within one. What this study adds to the band is therefore its local figure with declared uncertainty, from which an operational reference follows, since a 30% replacement is defensible when compression governs the element, whereas in elements governed by flexure the prudent threshold falls to 29%. Both values belong to the principal analysis and shift with the sources retained and the weighting applied (Section 2.6).

Read together, the four results sustain what no available Peruvian study could sustain on its own: that a replacement range with bounded loss exists, that this range can be read only in the set of studies and not in any one of them, and that each property has its own threshold under an explicit allowable strength loss. The contribution lies in the evidence on which those thresholds rest. The ten studies assembled here are Peruvian theses, i.e., the gray literature deposited in institutional repositories and absent from the indexes on which the available syntheses were built, and each was tested at one design strength against a control of its own; brought together in a database traceable to each laboratory value, they yield a decision parameter with its uncertainty declared, without testing a single new specimen. Normalization against the control of each study, an established step on which this synthesis rests, removes the differences in absolute strength between sources, so that the same synthesis can be repeated in any country with evidence equally scattered, where it yields a threshold of its own, estimated from the aggregates and the strength class of that

evidence. Each application requires, however, a minimum mass of sources per subgroup, and the results of this work make it possible to specify what happens when that mass is not reached.

3.2. Limitations and Future Research

The survey of the scope begins with the design, which determines the class of question this work answers. Assembling measurements already produced and expressing them on a common scale makes it possible to estimate a parameter of national scope that no isolated experiment delivers, and to do so without casting a single new specimen; in exchange, the evidence so built establishes with what magnitude replacement and strength move together, and not the mechanism by which they do so. The thresholds therefore operate as a reference for deciding rather than as a prediction of a particular test, and that same threshold opens the next step, since the study leaves a concrete hypothesis formulated where none existed before: to produce mixes at the ratios derived here, in a single laboratory and under a single curing protocol, and to verify whether the loss observed matches the loss estimated.

Such a test would moreover control what the assembled sources describe only in part, which is the quality of the recycled aggregate itself. In this design, that dimension receives an indirect and effective treatment, since each source is compared with itself before entering the model and the random intercept captures what distinguishes one study from another, so that the differences in adhered mortar and absorption that the literature identifies as decisive [11] are absorbed within the between-study variation, and their possible effect on the loss per point replaced was examined by allowing each study a slope of its own, which left the fit unchanged (Section 2.4). What that treatment does not separate is how much of that variation corresponds to aggregate quality, how much to the water–cement ratio and admixtures of each mix, and how much to the remaining differences in practice. In one of the sources, the mixes also differ in the use of admixture, so that replacement and admixture act together within that study. Setting that source aside keeps the slope negative and significant and raises the prudent threshold to 34% in compression and 33% in flexure (Supplementary S7), so the thresholds adopted remain the more prudent. Reference [20] showed that the separation is attainable once the datum is recorded, so that requiring the adhered mortar fraction and the absorption at extraction, and stratifying by them when the number of sources allows, is the route by which a replication would narrow the intervals and refine the threshold.

The material is not, however, the only thing that changes from one study to another, because the way of measuring it changes as well. The appraisal of the sources (Section 4.2) records that the nine sources of the principal analysis assessed flexure with third-point loading or without declaring the method, and at 28 days the third-point sources alone give a slope close to that of the complete set (Supplementary S7); its subgroups at 7 and 14 days, in turn, assemble four studies, a number with which the random component is estimated with less power, and the hypothesis testing was anchored at 28 days, the normative age of acceptance. Harmonizing the loading method before synthesizing, or stratifying by method at the early ages when the number of studies permits, would allow those ages to sustain a reading of the same standing as the one attained at 28.

From those two sources of variation, the material and the method, follows the scope with which the result must be read. The thresholds describe concrete of $f'c = 210$ kg/cm^2 made with the recycled aggregates of the Peruvian studies assembled, in its compressive and flexural strength. Other strength classes, other aggregates and properties such as durability and the modulus of elasticity have thresholds of their own, since the effect of replacement varies with the strength class [12] and with the adhered mortar and absorption of the aggregate [11]. Each of those extensions defines a study still to be done, and the

procedure is described so that another team can carry it out, since a replication needs only three conditions: an open-access repository with studies deposited in it, that each one report a control concrete without replacement, which is what makes normalization possible, and a set of studies sharing a design strength, so that the resulting threshold corresponds to one class and not to an average of classes.

Any replication depends, in the end, on where the evidence resides, and there the assembled set holds an advantage of origin: a thesis is deposited independently of the result obtained, so that a set built from them is less exposed to publication bias than one made up only of articles, where the result conditions appearance. That advantage does not amount, however, to a formal check. The eligibility criteria that make the sources comparable, a single design strength, a control concrete of their own, and complete test values, are also what determine how many satisfy them, and that number governs two things at once: bias tests demand more studies than so homogeneous a set admits, and between-study variance is estimated with an uncertainty that narrows the intervals below what they warrant. From this arises the research that follows from the present study: a cumulative synthesis of the Peruvian evidence that incorporates each thesis as it is deposited and recomputes the threshold at every update, applying the corrections based on the t distribution recommended when studies are few and their precisions differ [27], and submitting the set to formal bias testing as soon as it reaches the number that the estimation of a slope requires [28].

4. Materials and Methods

This study adopted a quantitative approach and was carried out as a systematized review with a quantitative synthesis of secondary data, aimed at determining and quantifying the extent to which the replacement ratio of natural coarse aggregate by RCA is associated with the mechanical performance of the concrete, in order to establish replacement thresholds with their associated uncertainty. No primary data were therefore generated by testing specimens in the laboratory; instead, the results reported by independent primary studies were compiled, cleaned, and statistically modeled in order to establish the aggregate trend in the mechanical performance of the material. The unit of analysis of this study is accordingly each mechanical strength result reported by a primary study, defined by the combination of the study of origin, the type of test, the curing age, and the replacement ratio evaluated.

4.1. Type and Design of the Study

The study is correlational in scope, in that it examines the association between a continuous independent variable, the replacement ratio of natural coarse aggregate by RCA obtained from concrete, and a dependent variable, the mechanical performance of the concrete, expressed in two dimensions that correspond to the specific hypotheses of the study: compressive strength and flexural strength, assessed at 7, 14, and 28 days. The design is non-experimental, since the production and testing conditions were not manipulated but come entirely from the original sources; the work was confined to observing, systematizing, and analyzing information that already existed. Because the sources consulted were published over a broad period, this study is moreover retrospective and longitudinal in its compilation, in that it integrates evidence produced at different points in time under a single analytical framework.

4.2. Document Population and Included Sources

The document population comprised Peruvian civil engineering professional-degree theses that assessed the mechanical performance of concrete produced with RCA obtained

from concrete at a design strength of f′c = 210 kg/cm^2, identified through a search of the National Open-Access Digital Repository of Science, Technology, and Innovation (ALICIA), the National Registry of Research Works (RENATI), and the institutional repositories of Peruvian universities, restricted to the period 2011 to 2025 and closed in May 2026. Although this body of documents belongs to the academic gray literature, its strength values are neither estimates nor opinions but laboratory test results reported in the documents themselves, which makes it the most extensive body of empirical evidence available on the behavior of the material with Peruvian aggregates, cements, and construction practices. Peruvian evidence published in indexed journals and proceedings at that same design strength was deliberately left outside the set, since it belongs to the document category whose absence from syntheses motivates this work, and it is used in the discussion as an external contrast to the results obtained.

Identification within those repositories relied on Spanish search terms tried in successive combinations, since their interfaces retrieve by free-term matching on the title, the abstract, and the keywords of each record. The productive terms belong to three families: the material, with concreto reciclado, agregado de concreto reciclado, agregado grueso reciclado, and agregado reciclado de demolición; the property assessed, with propiedades mecánicas and propiedades físicas y mecánicas del concreto; and the operation studied, with sustitución and reemplazo del agregado grueso natural. The material expressions acted as the anchor of every combination, since they are what distinguish this object within concrete research as a whole, while those of property and operation served to narrow the results the former returned. Supplementary S1 lists the full title of every source incorporated, in which those expressions reappear, so that the effective coverage of the strategy can be checked term by term.

The time restriction of that search answers to a requirement of comparability rather than to a preference for what is recent. Test methods, laboratory equipment, and curing procedures have become progressively more precise, so that assembling measurements separated by several decades would introduce into the set a source of variation foreign to the phenomenon under study, attributable to the technique with which the measurement was made and not to the material. The window was therefore set at the fifteen publication years between 2011 and 2025, a span that coincides moreover with the period of sustained expansion of research on concrete with recycled aggregate [22].

The incorporation of each thesis into that set was decided by explicit criteria, fixed before data extraction began and applied equally to all identified sources. Studies were considered relevant for the analysis when they used RCA obtained from demolition concrete as the replacement material, adopted a design strength of f′c = 210 kg/cm^2, included the control concrete without replacement as a reference, reported numerical values of compressive or flexural strength, and tested at one or more of the ages of 7, 14, or 28 days. Sources were discarded, by contrast, when they used replacement materials other than recycled concrete aggregate, such as PET plastic, tire rubber, polyvinyl chloride, glass, or steel fibers, as well as those presenting data duplicated from a source already incorporated, omitting the control concrete and thereby preventing subsequent normalization, or exhibiting values outside the expected physical range or signs of failure in the mix design of the control concrete.

Those criteria were applied to all the theses identified in the repositories consulted and by themselves delimited the set of eligible units. This study analyzed every one of them, so that no sampling stage intervened: instead of selecting a subset within the sources that met the requirements, the work proceeded with the entirety of the accessible evidence that satisfies the definition of the study. That set comprised ten primary studies. In nine of them, the recycled aggregate replaced part of the natural coarse aggregate, and

the 179 observations extracted from them form the analytical database of the principal analysis (Supplementary S3); in the remaining one, it entered the mix as an addition proportional to the cement mass, and its 24 observations are described separately (Section 2.6; Supplementary S1). Those ten sources were published between 2019 and 2024, so that the time criterion did not operate on them as an effective filter: the evidence satisfying the criteria is concentrated in the most recent period, and none were excluded by the lower bound of the window. The coverage thus achieved is bounded by the access route used, since only theses deposited in open-access repositories could be incorporated. The complete process of identification, screening, eligibility, and inclusion is summarized in Figure 3, and the 52 records identified are listed in Supplementary S1 with the stage at which each was retained or excluded and the reason for every exclusion.

With the set thus delimited, each of those ten sources was read in full text through document analysis, a technique that operated by means of a structured record sheet in a spreadsheet. For every result, it recorded the identification of the source, the type of test, the design strength, the test age, the replacement ratio, and the strength value reported, so that the ten studies were cast in a single format and their figures became comparable with one another. Because every record retains the source it comes from, any datum in the analysis can be traced back to the document that produced it; the complete catalog of those sources, with their persistent identifiers and their references, is presented in the supplementary material (Supplementary S1).

That same full-text reading served to appraise the methodological quality of each source, a check that is unavoidable when the evidence comes from the gray literature, since confidence in the synthesis rests on the tests that feed it having been carried out under standardized procedure. The appraisal drew on four features verifiable in the document itself: the test standard applied, the reporting of the mix design, the number of specimens tested, and the inclusion of laboratory certificates as an appendix, as set out in Table 9.

That fourth criterion separates the thesis from the published article and explains why this class of source is advantageous for a synthesis. Six of the ten sources append the calibration certificate of the testing machine and of the instruments used, with its certificate number; two include the cement quality certificate, one of them alongside the calibration certificate, and one the certificate of a bond test, so that eight of the ten are accompanied by at least one laboratory document, the remaining two being identified in Table 9. A journal article states the standard under which it measured and there its documentary chain ends, because editorial length allows no more, whereas a thesis of between 123 and 222 pages preserves the document attesting to the instrument with which the measurement was made. The certificate attests to that instrument, so that traceability reaches the recorded value, while specimen preparation and curing are covered by the three preceding criteria. What this study takes from those sources are, accordingly, test values verifiable at their origin and not the conclusions their authors drew from them.

4.3. Data Extraction and Processing

With the set of sources established, the work ceased to bear on the documents and came to bear on the data they contain. The strength values reported were extracted from each thesis and tabulated in long format, that is, a structure in which every row corresponds to a unique observation, determined by the combination of the study of origin, the type of test, the curing age, and the replacement ratio evaluated. That arrangement is what allows results produced under different experimental designs to coexist in a single matrix and to be treated as observations of one and the same phenomenon.

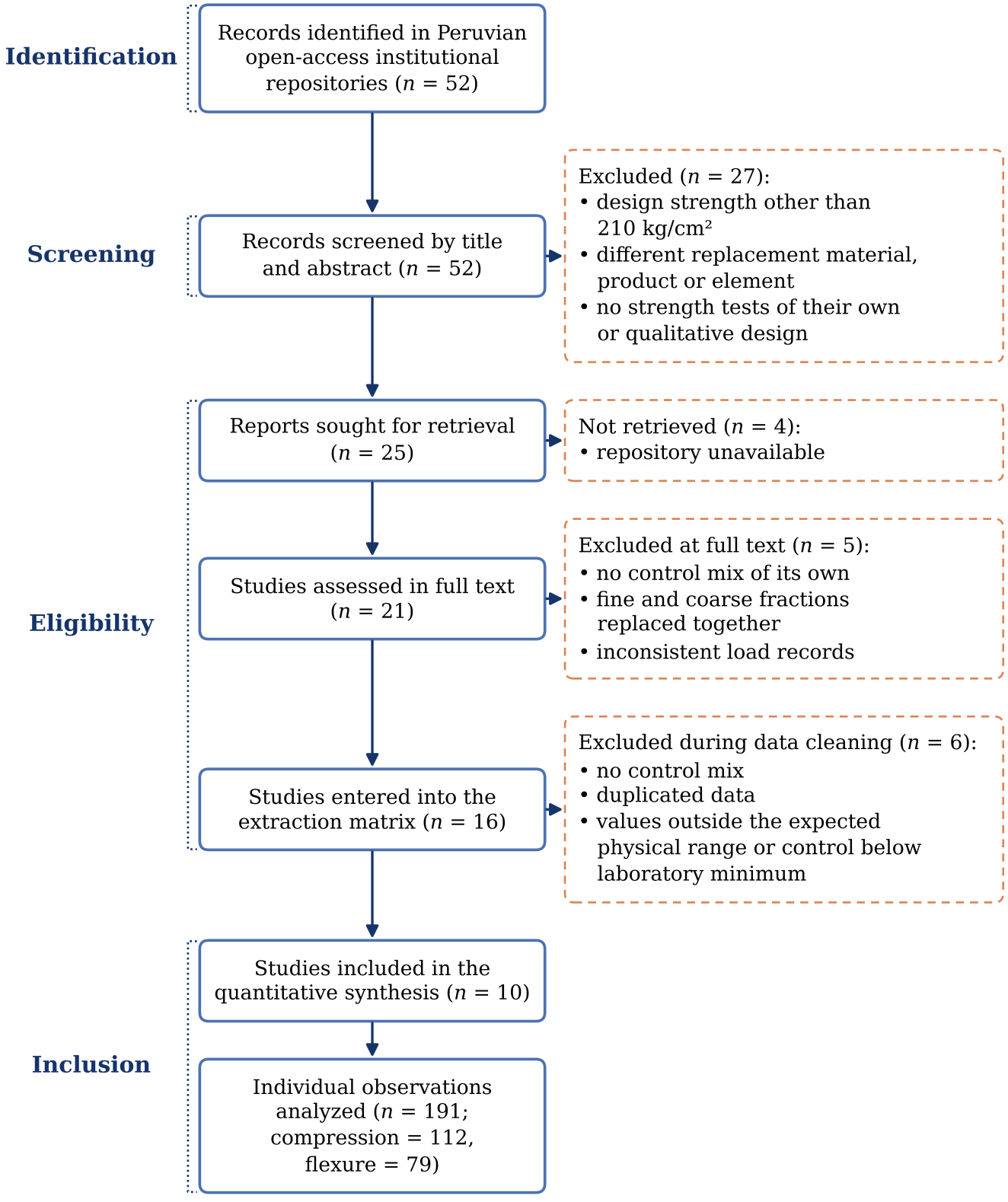


Figure 3. Flow diagram of the identification, screening, eligibility, and inclusion of the primary sources. Note: Diagram adapted from the PRISMA 2020 reporting logic [29] to a systematized review of the gray literature. Eligibility is kept as a stage of its own because it is where the full text was assessed and the extraction matrix was built; *n* = number of records or studies at each stage.

Table 9. Methodological quality appraisal of the included primary sources.

Label	Test Standards (Compression; Flexure)	Mix Design	Observations Contributed	Attached Certificates
Inv. 1	NTP 339.034/ASTM C39; ASTM C78	Yes (ACI 211)	24 (12 compression; 12 flexure)	Equipment calibration
Inv. 2	NTP 339.034/ASTM C39; NTP 339.078/ASTM C78	Yes (ACI 211)	16 (12 compression; 4 flexure)	Cement quality
Inv. 3	NTP 339.034/ASTM C39; NTP 339.078	Yes (ACI 211)	16 (12 compression; 4 flexure)	Equipment calibration
Inv. 4	ASTM C39; ASTM C293	Yes (ACI 211)	5 (5 compression)	Bond test
Inv. 5	NTP 339.034; NTP 339.078	Yes (ACI 211)	16 (12 compression; 4 flexure)	None
Inv. 9	Not reported	Yes (ACI 211)	30 (15 compression; 15 flexure)	Equipment calibration
Inv. 10	ASTM C39/NTP 339.034; ASTM C78/NTP 339.078	Yes (ACI 211.1)	12 (8 compression; 4 flexure)	Calibration and cement
Inv. 11	NTP 339.034/ASTM C39; NTP 339.078	Yes (ACI method)	30 (15 compression; 15 flexure)	Equipment calibration
Inv. 14	Not reported; center-point loading described	Not reported	24 (12 compression; 12 flexure), described separately	None
Inv. 15	Generic (ASTM/NTP, no specific designation)	Yes (ACI 211)	30 (15 compression; 15 flexure)	Equipment calibration

Note: The appraisal was carried out on the full text of each thesis (1 August 2026). Label denotes the source identifier used in the compilation matrix and in Supplementary S1, whose numbering is not consecutive because it preserves the original identifiers. Not reported indicates that the document does not cite the designation of the standard for the compression or flexure test, even though it describes the procedure. The observations column records the strength values that each source contributes to this synthesis, one for every combination of replacement ratio and test age, and their sum equals the 179 observations of the principal analysis plus the 24 of the source described separately; it should not be confused with the number of specimens each source tested, which is higher in every case, since each value comes from the average of several specimens and several sources also assessed design strengths or properties outside the scope of this study. Attached certificates indicate which laboratory document accompanies the thesis as a formal appendix. Inclusion followed the eligibility criteria, and this appraisal defines the sources set aside in the sensitivity analysis by documented quality, reported in the Results (Supplementary S7).

Cleaning acted on those observations and not on the sources, a distinction that matters because a thesis already incorporated could contribute defective records without ceasing to satisfy the eligibility criteria. At that level, duplicate records, outliers, records with scale inconsistencies, and records showing failure in the mix design were removed; the most illustrative case is that of an included source whose flexure observations were removed in their entirety, because they were measured on beams reinforced with a steel bar and therefore express the capacity of a reinforced element, while its compression observations, obtained on plain cylinders of the same mix and curing, were kept intact. Each of those decisions was recorded with its justification and the page of the source that supports it in the cleaning log (Supplementary S2), so that the passage from the original values to the analytical database can be reconstructed record by record. Each of the 203 records is, in turn, linked to its original record, to the page and table of its source, and to the number of specimens and their standard deviation where the source reports them (Supplementary S8). Gaps received a treatment different from that of defects: combinations of ratio and age

that a source did not test generated no observation and were left as such, without any imputation, since the absence of a test is not a missing datum but a condition that study did not evaluate.

With the database cleaned, the obstacle that prevents a direct comparison of results from independent studies still stood: each study starts from a control strength of its own, which varies with the quarry, the water–cement ratio, the type of cement, and the curing conditions, so that one and the same loss attributable to replacement is expressed in different absolute values depending on the study it comes from. To neutralize that heterogeneity, every value was transformed into relative strength, defined as the ratio between the strength obtained at a given replacement ratio and the strength of the control concrete of the same study, the same property, and the same test age, expressed as a percentage:

$$\mathrm{Rrel} = (\mathrm{fx}/\mathrm{f0}) \times 100$$

In this expression, fx denotes the strength reported at a replacement ratio x and f0 the strength of the control concrete, without replacement, of that same source, property, and age. With every observation referred in this way to its own control, concrete without recycled aggregate equals 100% in all sources and the ten are placed on a common scale, which is the one used in the analysis and the reason why the results in Section 2 are read as percentages. Those control values remain in the analysis as the point at 0% replacement in each study, from which its loss is measured, and the estimates at 28 days were repeated without them as a sensitivity analysis (Supplementary S4.7 and S4.8). Every group in a study is divided by that same control, so an error in it shifts all of them together, a shift that the random intercept of each study represents approximately, while each control value enters the model as a fixed reference, without its own measurement uncertainty. Each relative strength is the mean of the specimens tested by its source, between one and six, with three in 135 of the 167 observations whose source reports that number, and a further sensitivity analysis gave each observation a weight proportional to it (Supplementary S6), assigning the modal value of three to the remaining twelve. That weighting assumes that the specimens of every group share the same dispersion, so that the variance in each mean depends only on the number of specimens. The resulting analytical database, together with its variable dictionary, is provided as supplementary material (Supplementary S3).

4.4. Methods of Data Analysis

The analysis was organized on that common scale, advancing from the descriptive to the inferential level, and was run in jamovi, version 2.7.15 [30], which operates on the R language, version 4.5 [31]. The analysis files are made available as supplementary material, one for each subgroup, so that every model can be reproduced separately (Supplementary S4.1 to S4.6, with the sensitivity analyses in S4.7 to S4.12). Before any calculation, the data were segmented by type of test and by age, because compression and flexure respond to different failure mechanisms and because concrete at 7 days has not yet developed the strength it reaches at 28: treating those conditions together would have averaged behaviors that the study sets out to distinguish. The composition of the six subgroups so defined is shown in Table 1.

With that coverage in view, each subgroup was characterized by measures of central tendency and the dispersion of relative strength, which establish the magnitude and the variability of the phenomenon before it is put to the test.

The test with which to estimate the association depended on the shape of those data, and for that reason the distribution of relative strength was examined in each subgroup by the Shapiro–Wilk normality test. Its result decided the route: when normality does not hold and the expected relationship between replacement and strength is monotonic rather

than necessarily linear, the Spearman correlation coefficient is the one that answers to both conditions [32], and it was the one used to estimate the strength and the direction of the association in each of the six subgroups.

Correlation, however, treats every observation as if it were independent of the rest, and here it is not: the observations come from studies, and those of a single study share quarry, cement, curing, and operator. That dependence among observations nested within studies calls for a model that represents it explicitly [33], so that a linear mixed-effects model was fitted with the GAMLj module, version 3.6.5 [34], in which relative strength was the dependent variable, the replacement ratio entered as a fixed effect, and the study of origin as a random intercept. That intercept absorbs the level differences between studies and consequently returns more conservative standard errors than those of an analysis assuming full independence. The standard errors, confidence intervals, and probability values of the fixed effects were obtained with the parameters package [35], using the t distribution with Satterthwaite degrees of freedom, and at 28 days they were checked against Kenward–Roger degrees of freedom and a parametric bootstrap of 1000 replicates, both suited to a small number of groups (Supplementary S4.3 and S4.6).

That same intercept defines where the inference points. What the model estimates is the behavior of the material and not a property of the Peruvian body of documents: each study functions as one realization of the experimental conditions under which recycled aggregate may be used, and the variance in the intercept measures how far the result shifts on moving from one realization to another. The significance tests therefore report on the stability of the effect across conditions, which is the question the study poses. The proportion of total variability corresponding to those differences between studies is expressed by the intraclass correlation coefficient [36]. The fit of each model is described by the marginal and conditional R^2, the first for the variance explained by the replacement ratio and the second for that explained by the replacement ratio and the study together [37], while a likelihood-ratio test compares each model with the same model without the random intercept, an ordinary least squares regression, to establish whether the differences between studies justify including it. At 28 days, the same test also compared each model with one that allows each study a slope of its own (Supplementary S4.3 and S4.6), and the prudent criterion was also computed with the source that added recycled aggregate in proportion to the cement mass (Supplementary S4.9 and S4.10). In flexure, each source broke its control and its recycled mixes under the same loading configuration, so normalization against its own control removes the difference in level between loading configurations; at 28 days, the loading method, third-point or not declared, was also entered as a moderator of the slope, and the model was re-estimated with the third-point sources alone (Supplementary S7).

The precision with which that variability is estimated depends on the number of studies available, and in this set that number is small. With few studies, between-study variance is estimated with substantial uncertainty and tends to be underestimated [38,39], so that intervals built on it turn out narrower than they should be [27]. The intervals reported here therefore operate as a lower bound on the real uncertainty, and the thresholds derived from their limit as a reference for deciding rather than as a guarantee. The caution extends to all six subgroups and is sharpest in flexure at 7 and 14 days, where the number of studies is the smallest in the set and where the estimate is moreover checked against its corresponding correlation. All tests, both those of correlation and those of the model, were evaluated at a significance level of 0.05, which makes the statistical decisions reported in Section 2 comparable with one another.

The derivation of the replacement thresholds rests on those estimates. Since the fixed effect expresses the change in relative strength per percentage point of replacement, the expected loss for a given ratio follows from the product of the two quantities, and the

replacement compatible with a defined allowable strength loss follows from the inverse operation, computed in both cases from the point estimate and from each limit of the confidence interval of the slope. Both operations assume that the loss grows in proportion to the replacement ratio, an assumption examined at 28 days by adding a quadratic term in that ratio to each model (Supplementary S4.3 and S4.6). That derivation adds no new analysis but translates arithmetically the estimates already obtained, and its full development is documented in the supplementary material (Supplementary S5).

4.5. Ethical Considerations

Because the study drew exclusively on publicly available secondary information and did not involve the participation of humans or animals, it did not require approval by an ethics committee for research with living beings. The study nevertheless observed strictly the principles of scientific integrity applicable to the handling of sources. In particular, the intellectual authorship of each primary study was respected through explicit citation and referencing, avoiding any form of plagiarism and acknowledging the origin of every datum used; the catalog of the sources and their complete references are presented in the supplementary material (Supplementary S1). Honesty in the treatment of the information was likewise ensured, in that the values were transcribed faithfully from their sources, no data were manipulated or fabricated, and the cleaning decisions were documented transparently to allow their verification. All the information came from open-access institutional repositories, so that its use for academic purposes infringed no rights of confidentiality or ownership. Finally, in keeping with the principles of open science and reproducibility, the cleaned database, the complete catalog of the sources, and the analysis files are made available as supplementary material (Supplementary S1 to Supplementary S9).

4.6. Use of Generative Artificial Intelligence Tools

ChatGPT Classic (OpenAI, version 1.2026.184) and Claude for Mac (Anthropic, version 1.32885.1) assisted in drafting and editing the text and in writing and running the analysis and figure code; Scopus AI (Elsevier) supported the search and preliminary screening of the literature, and ChatGPT (OpenAI) rendered the graphical abstract from a description prepared by the authors. The authors verified every output against the database and the full text of the sources.

5. Conclusions

The replacement ratio of natural coarse aggregate by RCA is negatively associated with the mechanical performance of concrete of $f'c = 210$ kg/cm^2, with a loss close to a quarter of a point of relative strength per point replaced at 28 days. That magnitude, estimated after discounting the differences between studies, which account for between 23% and 51% of the observed variability, translates into a maximum replacement of 34% in compression and 29% in flexure under the allowable strength loss adopted. Both values belong to the principal analysis and shift with the sources retained and the weighting applied. Brought to the scale of the construction site, a replacement of 30% entails a mean loss of 7.05 percentage points in compression and 7.68 in flexure, which rise to 8.94 and 10.26 under the prudent criterion, the confidence-interval endpoint corresponding to the greater strength loss. Read on that basis, the ratio remains within the admitted loss when compression governs the element and exceeds it when flexure governs. These readings describe the mean behavior of the set of studies under the criterion adopted here, and neither bounds the loss of an individual mixture. Each value behind those figures can be traced to the page of its source thesis (Supplementary S8), so that the threshold can be checked and updated as new theses are deposited.

In compressive strength, the loss reaches 0.235 percentage points per point replaced, and the slope of the model is significant at each of the three ages. The result is read at 28 days because that is the age of acceptance and the one that gathers the most studies in both properties.

In flexural strength, the loss reaches 0.256 percentage points per point replaced, of similar magnitude to that of compression, and its estimate is less precise, with a confidence interval about 40% wider. That width, associated with fewer studies and greater dispersion, together with the slightly steeper slope, is what lowers the prudent threshold in flexure to 29%, so that the threshold for concrete with recycled aggregate is set by the property whose evidence is least precise.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/recycling11120213/s1: Supplementary S1: Catalog of primary sources and list of the records screened; Supplementary S2: Database, with the dataset to which each record belongs; Supplementary S3: Analytical database of the principal analysis (179 observations from nine sources); Supplementary S4.1 to S4.6: jamovi files of the principal analysis for compression and flexure at 7, 14, and 28 days; Supplementary S4.7 and S4.8: jamovi sensitivity analyses at 28 days without the controls; Supplementary S4.9 and S4.10: jamovi sensitivity analyses at 28 days with the source that added recycled aggregate; Supplementary S4.11 and S4.12: jamovi sensitivity analyses at 28 days with the records left out of the database; Supplementary S5: Complete statistics of the principal analysis; Supplementary S6: Sensitivity analyses weighted by specimen count and by the reported standard deviations (R script); Supplementary S7: Sensitivity analyses by documented source quality and flexural loading method (R script); Supplementary S8: Traceability registry of the 203 records; Supplementary S9: Map of the datasets used in each analysis; Supplementary S9.1: Datasets and key figures with their source files; Supplementary S9.2: R script that reproduces the principal and sensitivity results.

Author Contributions: S.S.P.: data curation, formal analysis, investigation, project administration, resources, software, writing—original draft. J.I.B.: data curation, formal analysis, investigation, project administration, resources, software, writing—original draft. E.J.M.-S.: conceptualization, data curation, formal analysis, investigation, methodology, supervision, validation, visualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The relevance of ethical review was assessed through the applicable institutional procedure. The study used exclusively publicly available secondary data, so that review by an ethics committee for research with living beings did not apply.

Informed Consent Statement: Not applicable.

Data Availability Statement: The cleaned database, the complete catalog of the primary sources, the analyzed matrix, and the statistical output are made available as supplementary material (Supplementary S1 to Supplementary S9) and are deposited in the Zenodo public repository under a CC BY 4.0 license at https://doi.org/10.5281/zenodo.22051012 [40]. Any additional information is available on request from the corresponding author.

Acknowledgments: During the preparation of this study, the authors used ChatGPT Classic (OpenAI, version 1.2026.184), Claude for Mac (Anthropic, version 1.32885.1), Scopus AI (Elsevier), and ChatGPT (OpenAI) for the purposes described in Section 4.6. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pacheco, J.; de Brito, J. Recycled aggregates produced from construction and demolition waste for structural concrete: Constituents, properties and production. *Materials* **2021**, *14*, 5748. [CrossRef] [PubMed]
2. Watari, T.; Cao, Z.; Serrenho, A.C.; Cullen, J. Growing role of concrete in sand and climate crises. *iScience* **2023**, *26*, 106782. [CrossRef] [PubMed]
3. Lu, L. Optimal replacement ratio of recycled concrete aggregate balancing mechanical performance with sustainability: A review. *Buildings* **2024**, *14*, 2204. [CrossRef]
4. Soto-Paz, J.; Arroyo, O.; Torres-Guevara, L.E.; Parra-Orobio, B.A.; Casallas-Ojeda, M. The circular economy in the construction and demolition waste management: A comparative analysis in emerging and developed countries. *J. Build. Eng.* **2023**, *78*, 107724. [CrossRef]
5. Thives, L.P.; Ghisi, E.; Thives Júnior, J.J. An outlook on the management of construction and demolition waste in Brazil. *Clean. Mater.* **2022**, *6*, 100153. [CrossRef]
6. Bravo, J.; Valderrama, C.; Ossio, F. Cuantificación económica de los residuos de construcción de una edificación en altura: Un caso de estudio. *Inf. Tecnol.* **2019**, *30*, 85–94. [CrossRef]
7. Valenzuela Dorado, J.D.; Sacanambuy Urbano, D.A.; Duarte Goyes, A.; Cuellar Tonguino, M. Diagnóstico del manejo de los residuos de construcción y demolición, en la zona urbana del municipio de Sibundoy (Putumayo). *Luna Azul* **2025**, *61*, 259–298. [CrossRef]
8. Ziegler-Rodríguez, K.; Margallo, M.; Aldaco, R.; Irabien, A.; Vázquez-Rowe, I.; Kahhat, R. Environmental performance of Peruvian waste management systems under a life cycle approach. *Chem. Eng. Trans.* **2018**, *70*, 1753–1758. [CrossRef]
9. Abarca, Z.; Barreto, U.; Abarca, Y.; Barreto, O. Gestión de residuos de construcción y demolición en entornos urbanos desde la calidad ambiental. *Rev. Venez. Gerenc.* **2024**, *29*, 778–792. [CrossRef]
10. Medrano-Sánchez, E.J.; Fabian, M.C. Challenges and solutions in reinforcing self-built homes: Experiences from Villa María del Triunfo, Lima, Peru. *Front. Built Environ.* **2025**, *11*, 1711028. [CrossRef]
11. Muhammad, F.; Harun, M.; Ahmed, A.; Kabir, N.; Khalid, H.R.; Hanif, A. Influence of bonded mortar on recycled aggregate concrete properties: A review. *Constr. Build. Mater.* **2024**, *432*, 136564. [CrossRef]
12. Chang, Y.C.; Wang, Y.Y.; Zhang, H.; Chen, J.; Geng, Y. Different influence of replacement ratio of recycled aggregate on uniaxial stress-strain relationship for recycled concrete with different concrete strengths. *Structures* **2022**, *42*, 284–308. [CrossRef]
13. Vu, X.H.; Vo, T.C.; Phan, V.T. Study of the compressive strength of concrete with partial replacement of recycled coarse aggregates. *Eng. Technol. Appl. Sci. Res.* **2021**, *11*, 7191–7194. [CrossRef]
14. Makul, N.; Fediuk, R.; Amran, M.; Zeyad, A.M.; Klyuev, S.; Chulkova, I.; Ozbakkaloglu, T.; Vatin, N.; Karelina, M.; Azevedo, A. Design strategy for recycled aggregate concrete: A review of status and future perspectives. *Crystals* **2021**, *11*, 695. [CrossRef]
15. Butler, L.; Tighe, S.; West, J. Guidelines for selection and use of coarse recycled-concrete aggregates in structural concrete. *Transp. Res. Rec.* **2013**, *2335*, 3–12. [CrossRef]
16. Castillo Marín, A.J.; Tacilla Culqui, C.Y.; Alva Sarmiento, A. Use of recycled coarse aggregate in concrete f'c = 210 kg/cm^2 when incorporating percentages of recycled coarse aggregate. In Proceedings of the 22nd LACCEI International Multi-Conference for Engineering, Education, and Technology, San José, Costa Rica, 17–19 July 2024. [CrossRef]
17. Machaca Mamani, E.B.; Garcia Castro, R.A. Mechanical performance and structural feasibility of recycled aggregate concrete for sustainable construction. *Int. J. Adv. Technol. Eng. Explor.* **2025**, *12*, 1175–1189. [CrossRef]
18. Alva Pajares, L.D.; Salazar Idrogo, L.E.; Aguilar Aliaga, O. Compressive strength of concrete f'c = 210 kg/cm^2 with coarse aggregate substitution at 20%, 25%, and 30% using recycled concrete, Cajamarca 2023. In Proceedings of the 23rd LACCEI International Multi-Conference for Engineering, Education, and Technology, Mexico City, Mexico, 16–18 July 2025. [CrossRef]
19. Medrano-Sánchez, E.J.; Alberco, M.D.; Martos, E. Concrete enriched with prickly pear cactus ash: A leap towards sustainable construction. *Chem. Eng. Trans.* **2025**, *119*, 523–528. [CrossRef]
20. Silva, R.V.; de Brito, J.; Dhir, R.K. The influence of the use of recycled aggregates on the compressive strength of concrete: A review. *Eur. J. Environ. Civ. Eng.* **2015**, *19*, 825–849. [CrossRef]
21. Sivayogaraj, A.; Elavenil, S. Municipal solid waste incineration bottom ash in concrete: A systematic review and meta-analysis. *Sci. Rep.* **2025**, *15*, 39931. [CrossRef] [PubMed]
22. Cornejo-Meza, M.; Rubio-Cunishpuma, G.; Escobar-Segovia, K.; Garcia-Troncoso, N. Recycled aggregate concrete research (1983–2025): A global bibliometric and thematic evolution analysis for sustainable material design. *Buildings* **2026**, *16*, 1898. [CrossRef]
23. Zega, C.J.; Falcone, D.D.; Marfil, S.; Coelho dos Santos, G.S.; Di Maio, A.A. Residual alkali–silica reaction of coarse recycled concrete aggregate based on the application of a prescriptive measure from the Argentinian regulation. *Eur. J. Environ. Civ. Eng.* **2024**, *28*, 573–590. [CrossRef]
24. Pico-Molineros, A.; Sanchez-Rojas, E.; Bompa, D.V.; García-Troncoso, N. Incorporation of recycled concrete and mixed recycled aggregate as coarse aggregate replacements in structural grade concrete. *Constr. Build. Mater.* **2026**, *511*, 145303. [CrossRef]

25. Godin, K.; Stapleton, J.; Kirkpatrick, S.I.; Hanning, R.M.; Leatherdale, S.T. Applying systematic review search methods to the grey literature: A case study examining guidelines for school-based breakfast programs in Canada. *Syst. Rev.* **2015**, *4*, 138. [CrossRef] [PubMed]
26. Ministerio de Vivienda, Construcción y Saneamiento. *Norma Técnica E.060 Concreto Armado*; Reglamento Nacional de Edificaciones. Decreto Supremo N.º 010-2009-VIVIENDA. Technical Report; Diario Oficial El Peruano: Lima, Peru, 2009.
27. Röver, C.; Knapp, G.; Friede, T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. *BMC Med. Res. Methodol.* **2015**, *15*, 99. [CrossRef] [PubMed]
28. Jenkins, D.G.; Quintana-Ascencio, P.F. A solution to minimum sample size for regressions. *PLoS ONE* **2020**, *15*, e0229345. [CrossRef] [PubMed]
29. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
30. The Jamovi Project. *Jamovi (Version 2.7.15) [Computer Software]*; The Jamovi Project: Sydney, Australia, 2025. Available online: https://www.jamovi.org (accessed on 10 September 2026).
31. R Core Team. *R: A Language and Environment for Statistical Computing (Version 4.5) [Computer Software]*; R Foundation for Statistical Computing: Vienna, Austria, 2025. Available online: https://cran.r-project.org (accessed on 10 September 2026).
32. Navarro, D.J.; Foxcroft, D.R. *Learning Statistics with Jamovi: A Tutorial for Beginners in Statistical Analysis*; Open Book Publishers: Cambridge, UK, 2025. [CrossRef]
33. Hox, J.J.; Moerbeek, M.; van de Schoot, R. *Multilevel Analysis: Techniques and Applications*, 3rd ed.; Routledge: New York, NY, USA, 2018.
34. Gallucci, M. *GAMLj: General Analyses for Linear Models (Version 3.6.5) [Jamovi Module]*; University of Milano-Bicocca: Milan, Italy, 2019. Available online: https://gamlj.github.io/ (accessed on 10 September 2026).
35. Lüdecke, D.; Ben-Shachar, M.S.; Patil, I.; Makowski, D. Extracting, computing and exploring the parameters of statistical models using R. *J. Open Source Softw.* **2020**, *5*, 2445. [CrossRef]
36. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. Erratum in *J. Chiropr. Med.* **2017**, *16*, 346. https://doi.org/10.1016/j.jcm.2017.10.001. [CrossRef] [PubMed]
37. Nakagawa, S.; Schielzeth, H. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **2013**, *4*, 133–142. [CrossRef]
38. Langan, D.; Higgins, J.P.T.; Simmonds, M. Comparative performance of heterogeneity variance estimators in meta-analysis: A review of simulation studies. *Res. Synth. Methods* **2017**, *8*, 181–198. [CrossRef] [PubMed]
39. Iaquinto, S.; Bührer, L.; Feldmann, M.; Latal, B.; Held, U. How to quantify between-study heterogeneity in single-arm evidence synthesis?—It depends! *Syst. Rev.* **2025**, *14*, 138. [CrossRef] [PubMed]
40. Pizan, S.S.; Bendezú, J.I.; Medrano-Sanchez, E.J. *Database of Compressive and Flexural Strength of Concrete Containing Recycled Coarse Aggregate at 20.6 MPa (210 kg/cm^2) Design Strength: Ten Peruvian Primary Studies, 203 Records (179 in the Principal Analysis)*; Zenodo: Geneva, Switzerland, 2026. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.