Table S1. Information about the on-site school measurements.

Number	School name	Location	Number of buildings	Number of floors	Building condition	Offices	Toilets
1	Community Primary School	Ajede , Ijebu East	3	1	1 building under construction, 1 building to be renovated and toilet to be renovated	1	1
2	Community Primary School	Onipetesi, Ijebu East	5	1	3 buildings to be totally renovated, 1 building in habitable condition, toilet to be	1	1
3	Community Primary School	Orita, Ijebu East	9	1	1 building is habitable,3 buildings to be totally renovated, 1 building to be partially renovated, 1 newly constructed toilet, 3 old toilets	3	4
4	Moslem Primary School	Itapanpa, Ijebu East	3	1	1 dilapidated building ton be totally renovated, 2 habitable building, 2 toilets to be reconstructed	1	2
5	Moslem Primary School	Itapanpa, Ijebu East	3	1	1 habitable building, 1 building to be re- roofed, 1 building to be renovated	1	0
6	St Anthony Catholic Primary School	Terelu Imobi, Ijebu East	3	1	1 habitable building, 2 dilapidated buildings to be totally renovated	1	0
7	St. Benedict Catholic Primary School	Itasin, Ijebu East	3	1	2 habitable buildings, 1 to be renovated	1	0
8	St James Primary School,	Fotedo Imobi, Ijebu East	5	1	2 habitable buildings, 2 buildings to be renovated, toilet to be renovated	1	1
9	St. John Anglican Primary School	Itele, Ogun Ijebu East	4	1	1 building is under construction, 1 dilapidated building to be totally renovated, 2 habitable building	2	1
10	St John Catholic Primary School,	Ogbere, Ogun Ijebu East	4	1	2 habitable buildings, 2 buildings to be totally renovated	1	1
11	St. Luke Catholic Primary School	Lopo Korede, Ijebu East	3	1	2 habitable buildings, 1 building under renovation	0	0
12	Saint Mark R.C.M Primary School	Kajola, Ijebu East	1	1	1 building to be renovated	1	0
13	St Martins Catholic Primary School	Ebute Imobi, Ijebu East	3	1	1 habitable building, 2 buildings to be renovated	1	0
14	St Michael's Anglican Primary school	Imobi, Ijebu East	4	1	2 habitable buildings, 1 under construction, 1 abandoned building	1	0
15	St. Paul Anglican Primary School	Sasa Olumogo, Ijebu East	4	1	2 buildings and 2 toilets to be totally renovated	1	2
16	Community Primary School	Aberu, Ijebu East	3	1		0	0
17	Community primary school	Olooji, Ijebu East	3	1	1 dilapidated building, 1 habitable building and 1 building under construction	0	0

18	Community Primary School Idiegun Ademola	Idiegun Ademola, Ijebu East	1	1		0	0
19	Community Primary School	Ehin-Osun, Ijebu North East	3	1	1 abandoned building	1	0
20	St. Paul Anglican Primary School	Oguru , Ijebu East	3	1	1 newly built building	1	0
21	St Annes Catholic Primary School Irawo	Irawo Itamapako, Ijebu Ode	3	1		1	0

Table S2. On situ measurements compared to satellite images measurements. For the Google Earth images we have included uncertainty of 10%. In this analysis, we calculated uncertainty of 10% in the teaching area (after the veranda and offices are removed). The uncertainty associated with the external school area would be expected to be lower than this, and, indeed in all cases where the difference between the on situ and satellite measurements is much smaller than these uncertainties.

Number	School name	On situ measurement (total school buildings in m²)	Google Earth measurement (total school buildings in m²) (± 10%- uncertainty associated)
1	Community Primary School, Ajede , Ijebu East	427	429 (± 43)
2	Community Primary School, Onipetesi, Ijebu East	760	749 (±75)
3	Community Primary School , Orita, Ijebu East	956	949 (± 95)
4	Moslem Primary School, Itapanpa, Ijebu East	620	632 (± 63)
5	Moslem Primary School, Itapanpa, Ijebu East	530	542 (± ± 54)
6	St Anthony Catholic Primary School, Terelu Imobi, Ijebu East	702	715 (± 72)
7	St. Benedict Catholic Primary School, Itasin, Ijebu East	718	706 (± 71)
8	St James Primary School, Fotedo Imobi, Ijebu East	689	695 (± 70)
9	St. John Anglican Primary School, Itele, Ogun Ijebu East	866	876 (± 88)
10	St John Catholic Primary School, Ogbere, Ogun Ijebu East	745	742 (± 74)
11	St. Luke Catholic Primary School, Lopo Korede, Ijebu East	570	565 (± 57)
12	Saint Mark R.C.M Primary School, Kajola, Ijebu East	147	149 (± 15)
13	St Martin's Catholic Primary School, Ebute Imobi, Ijebu East	537	524 (± 52)
14	St Michael's Anglican Primary School, Imobi, Ijebu East	425	436 (± 44)
15	St. Paul Anglican Primary School, Sasa Olumogo, Ijebu East	418	426 (± 43)
16	Community Primary School, Aberu, Ijebu East	680	668 (± 67)
17	Community primary school, Olooji, Ijebu East	424	452 (± 45)
18	Community Primary School Idiegun Ademola, Idiegun Ademola, Ijebu East	206	215 (± 22)
19	Community Primary School, Ehin-Osun, Ijebu North East	467	452 (± 45)
20	St. Paul Anglican Primary School, Oguru, Ijebu East	576	561 (± 56)

21	St Anne's Catholic Primary School Irawo, Irawo Itamapako, Ijebu Ode	576	592 (± 59)

Table S3. Python codes used to calculate school overcrowding and the associated uncertainty.

Supplementary material on the software used to calculate school overcrowding and the associated uncertainty

The teaching area for each school was calculated using equation 1, presented in the main paper with terms defined in section 2.5.1.

$$A'_{\mathrm{T},j} = \left[\sum_{i=1}^{n} L_i (W_i - W_{\mathrm{V},i})\right] \times \left(1 - S_{\mathrm{Off}} B_j\right)$$

To apply this equation to each of the 1900 schools in the study set, software was written in Python 3 that used pandas for data handling, numpy for random number generation and matplotlib for graphing.

Monte Carlo Analysis required us to be able to run the software multiple times with, on each iteration, slight changes to the width and length of each building, and to the width of the veranda as well as to the fraction of building taken up by offices. For this, we established functions that allowed a random value to be added to a given width. For example, the routine which changed the width of the building looks like this:

```
def change_width(length : float, width : float, min_length_building : float, \
   min_width_building : float, uncertainty: float):
   """ If a building is too small it removes it. If it's not removed it adds an uncertainty to the width"""
   if is_it_blank(width) or is_it_blank(length):
      width=0.0
   elif width < min_width_building and length < min_length_building:
      width=0.0
   else:
      width = width + np.random.normal(0,uncertainty)
   return width
```

This particular routine uses the normal distribution for the random number generation, with a standard deviation equal to the uncertainty given in Table 4 (presented in the main manuscript). The function to change the length took the same form. And the veranda area was similar, but it calculated the full area (length times width) and not just one of those.

```
def calc_veranda_area(length : float, width : float, width_veranda_large : float, \
width_veranda_small : float, size_for_large_veranda : float, \
size_for_small_veranda : float, uncertainty_veranda : float):
   """ If a building is big enough to have verandas this will calculate a veranda area that can (later) be subtracted from that building's area. The uncertainty in the veranda width is used to create an error from a Gaussian distribution described by that uncertainty. Use uncertainty of zero to avoid randomness. """

   if width>=size_for_large_veranda:
      veranda_area=(width_veranda_large + np.random.normal(0,uncertainty_veranda))*length
   elif width>=size_for_small_veranda:
      veranda_area=(width_veranda_small + np.random.normal(0,uncertainty_veranda))*length
   else:
      veranda_area = 0
   return veranda_area
```

For the office area, on the other hand, we wanted a uniform distribution for office areas from 10% to 20%. That used the uniformly distributed random number generator. For this application, where the uncertainties are themselves uncertain, the randomness of the numpy random number generator was sufficient.

```
def calc_office_area(number_buildings : int, number_buildings_for_office : int, \
school_area_before : float, range_office_proportion : tuple):
   """ This calculates the total area in metres squared that is taken up by offices. If the school has too few buildings, we assume there is no office (school too small). Otherwise, we assume the office is some proportion between the low and high number of the tuple 'range office proportion' as a fraction of the total school area. To avoid randomness, make the two numbers in the tuple the same!"""
```

```
    low = range_office_proportion[0]
    high = range_office_proportion[1]

    if number_buildings < number_buildings_for_office:
        office_area = 0
    else:
        office_area = school_area_before * (np.random.uniform(low, high))
    return office_area
```

These routines for the width, length and veranda area were called as part of a broader function to calculate school teaching areas. This broader function could be run with a Boolean called "with_uncertainties" set to either True or False. If it was set as True, then the uncertainties in Table 4 were included. If it was set to False, then uncertainties were set to 0, so the routine did not make any change to building areas or veranda sizes.

These routines were used within the main program that used panda data frames and a csv file of all the school building lengths and widths to calculate equation 1, first for the original values (with_uncertainties = False) and then 50 times (in a for loop) with new random numbers (with_uncertainties = True). This provided, for each school, 1 teaching area for the original data set and 50 further teaching areas with those random numbers applied.

In supplementary material section S3, we show the results of those trials for two of the schools and explain why generally, the original data set value was in the middle of the range of 50 values from the Monte Carlo trials, but for a few schools (with a building close to our border for having/not having a veranda) it could be close to the top or bottom of that range.

Further basic statistical analysis and graphing were performed in Microsoft Excel to plot the ranges of values and to understand the uncertainties better (e.g. figures S2 and S3).

Table S4. Further explanation of differences between original calculations and the Monte Carlo analysis.

As an example, for both cases when the original calculation (before Monte Carlo) falls at the top or bottom (Figure S1), we selected school ID 6 (blue line in Figure 6) and school ID 36 (grey line in Figure S1), as both contain widths that are at the edge of the veranda conditions (Table 4 and equation 1). For example, school number 6 has two buildings that measure (W) 6.5 × (L) 22.5 m^2 and the other (W) 4.9 × (L) 22.3 m^2. The first two buildings have a width just on the edge of our condition on the veranda, thus for the original analysis (before Monte Carlo), the veranda is removed (1.6 m), decreasing the available area per pupil. The Monte Carlo simulation will add a positive error to this width about half the time. This will increase, slightly, the teaching area available for this building as the width will be wider and therefore the building bigger. But, for the other half of the time, the negative error will decrease the width of the building and therefore reduce it to below the veranda condition. As the veranda width of 1.6 m is much larger than the uncertainty on the width (0.5 m), this will also have the effect of increasing the teaching area available (slightly smaller school, but big veranda no longer removed). Thus, in almost all cases, the Monte Carlo simulation increased the teaching area of one of the two buildings compared to the original case.

On the other hand, school 36 shows the opposite situation which is also due to the veranda condition. This school has six buildings, with three of these having measurements widths of 6.4 m, 6.0 m, and 6.0 m respectively. These are all below the condition and so for the original case (black diamond), no veranda is removed from any of them. However, the Monte Carlo analysis can "pick" an error that moves any one of them above the criteria for having a veranda, and for the 6.4 m building, there is a high chance of that happening.

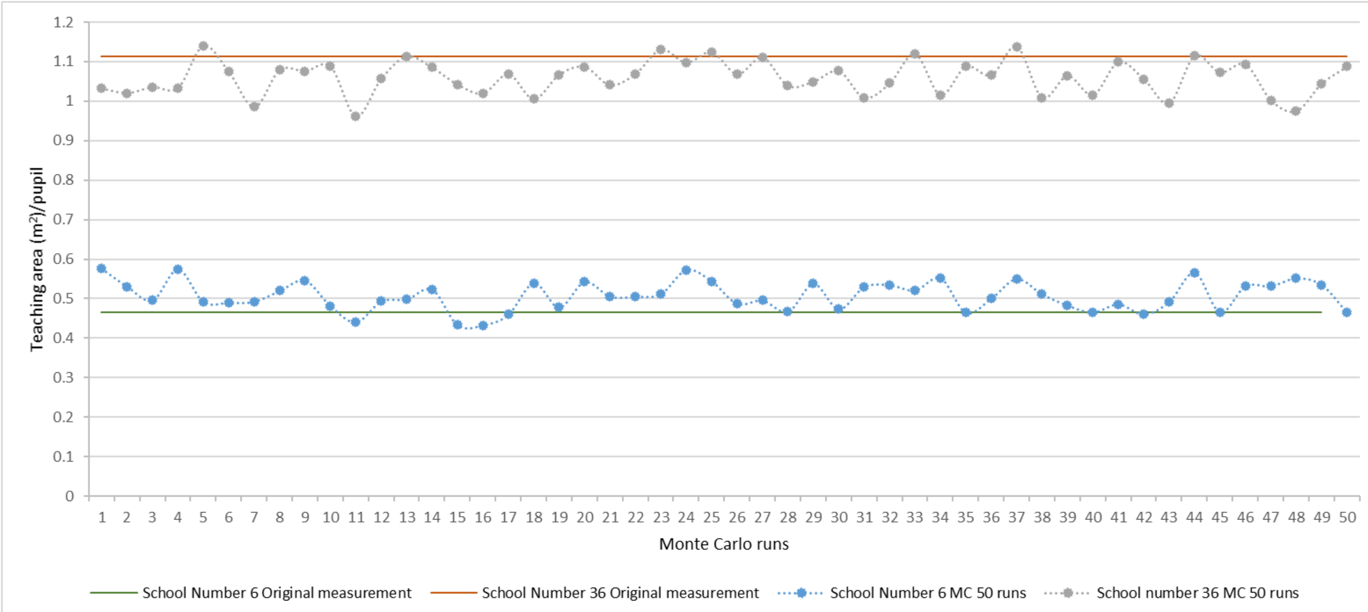

Figure S1. School number 6 (blue dots) and 36 (grey dots) showing the 50 Monte Carlo (MC) runs and the original measurements (brown and green lines) on the teaching area (m2)/pupil.

Table S5. Monte Carlo Analysis for individual schools.

To investigate both the bias (difference between the average of the Monte Carlo output and the originally determined value) and the standard deviation of the Monte Carlo output, we plot these as a function of school teaching area per pupil in Figure S2. There is a strong sensitivity to the calculated teaching area per pupil, and so these results are best described as relative values in S3. From these results, it is reasonable to consider the uncertainty associated with an individual school's "teaching area per pupil" to be around 10 %. If a more robust estimate of uncertainty is desired, then a model could be developed, with the uncertainty slightly larger than this for small teaching areas per pupil and slightly smaller than this for larger teaching areas per pupil.

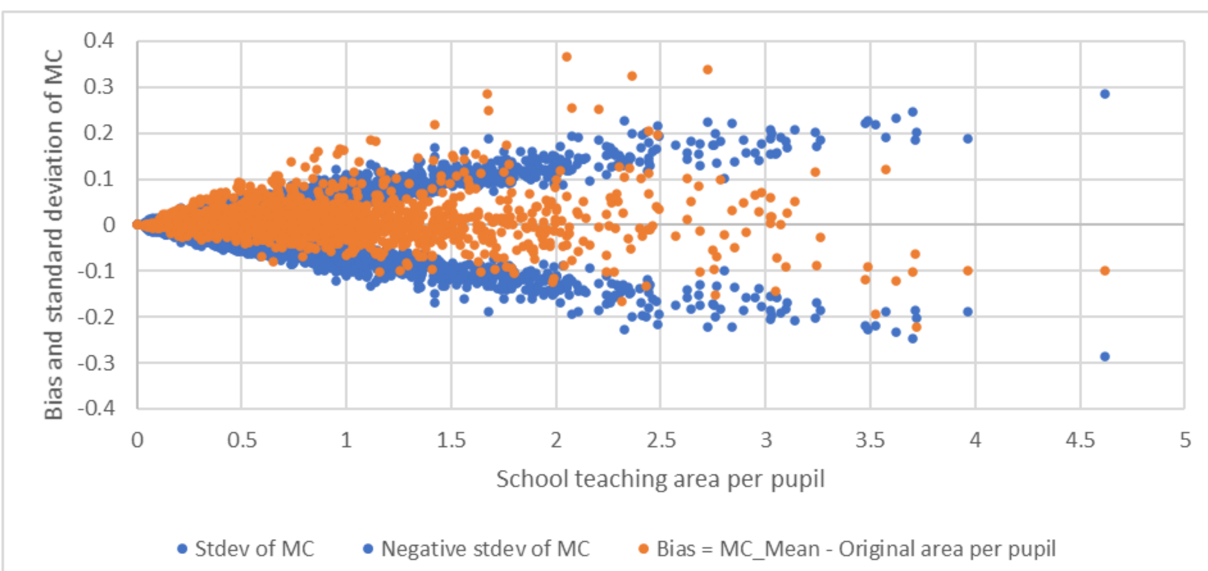

Figure S2. In blue is the standard deviation of the school teaching area per pupil as a function of school teaching area per pupil. Negative values are given (calculated as -1 times the standard deviation) as well, to show the full spread. In orange, the bias, calculated as the difference between the mean of the Monte Carlo output and the teaching area per pupil calculated from the original dataset.

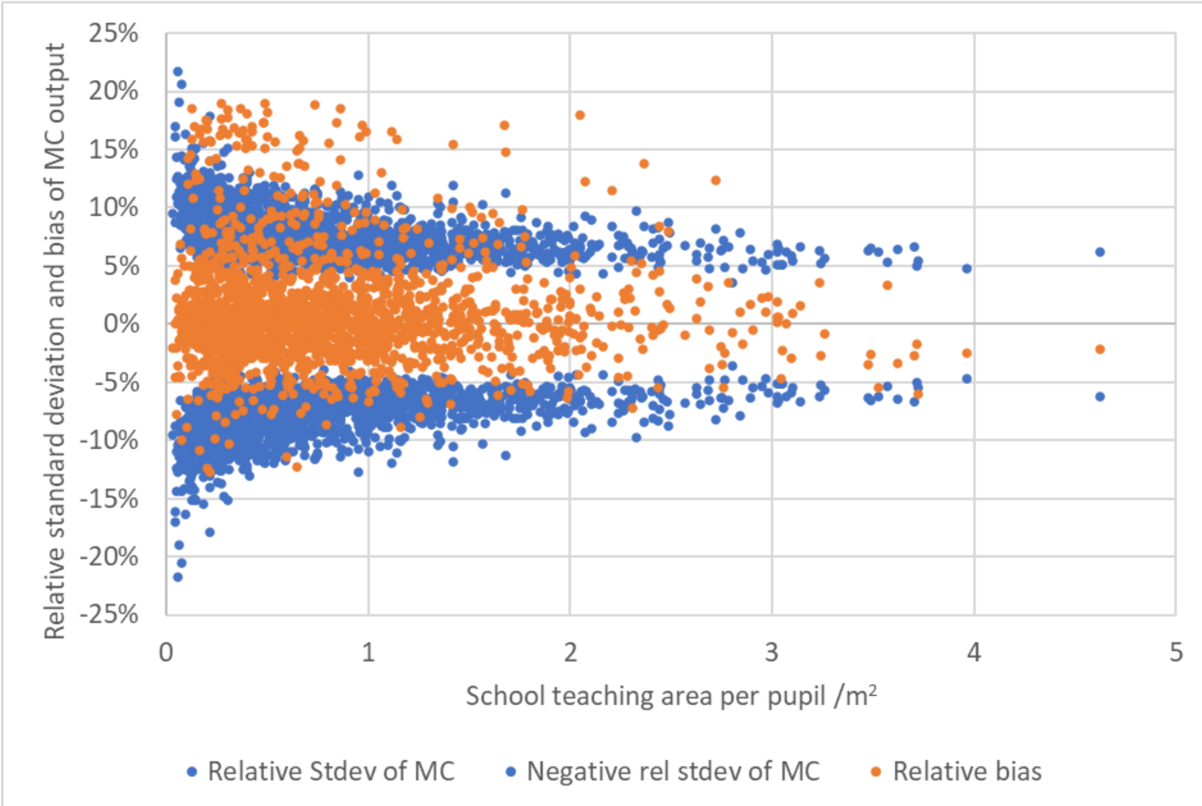

Figure S3. In blue is the standard deviation of the school teaching area per pupil as a function of school teaching area per pupil. Negative values are given (calculated as -1 times the standard deviation) as well, to show the full spread. In orange, the bias, calculated as the difference between the mean of the Monte Carlo output and the teaching area per pupil calculated from the original dataset.