

1 **Building knowledge of university campus population dynamics to enhance near-**
2 **to-source sewage surveillance for SARS-CoV-2 detection**

3 Chris Sweetapple^{a,b}, Peter Melville-Shreeve^b, Albert S. Chen^b, Jasmine M. S. Grimsley^a,
4 Joshua T. Bunce^{a,c,d}, William Gaze^e, Sean Fielding^f, Matthew J. Wade^{a,d,*}

5 *^a Joint Biosecurity Centre, Department of Health and Social Care, Windsor House, Victoria*
6 *Street, London, SW1H 0TL United Kingdom*

7 *^b Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences,*
8 *University of Exeter, Exeter, EX4 4QF United Kingdom*

9 *^c Department for Environment, Food and Rural Affairs, Seacole Building, London, SW1P*
10 *4DF United Kingdom*

11 *^d School of Engineering, Newcastle University, Newcastle-upon-Tyne, NE1 7RU United*
12 *Kingdom*

13 *^e Environment and Sustainability Institute, University of Exeter, Penryn Campus, Penryn,*
14 *Cornwall, TR10 9FE United Kingdom*

15 *^f Innovation Centre, University of Exeter, Exeter, EX4 4RN United Kingdom*

16 ** Corresponding author. Email: Matthew.Wade@dhsc.gov.uk*

17 **ABSTRACT**

18 Wastewater surveillance has been widely implemented for monitoring of SARS-CoV-2
19 during the global COVID-19 pandemic, and near-to-source monitoring is of particular
20 interest for outbreak management in discrete populations. However, variation in population
21 size poses a challenge to the triggering of public health interventions using wastewater

22 SARS-CoV-2 concentrations. This is especially important for near-to-source sites that are
NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

23 subject to significant daily variability in upstream populations. Focusing on a university
24 campus in England, this study investigates methods to account for variation in upstream
25 populations at a site with highly transient footfall and provides a better understanding of the
26 impact of variable populations on the SARS-CoV-2 trends provided by wastewater-based
27 epidemiology. The potential for complementary data to help direct response activities within
28 the near-to-source population is also explored, and potential concerns arising due to the
29 presence of heavily diluted samples during wet weather are addressed. Using wastewater
30 biomarkers, it is demonstrated that population normalisation can reveal significant
31 differences between days where SARS-CoV-2 concentrations are very similar. Confidence in
32 the trends identified is strongest when samples are collected during dry weather periods;
33 however, wet weather samples can still provide valuable information. It is also shown that
34 building-level occupancy estimates based on complementary data aid identification of
35 potential sources of SARS-CoV-2 and can enable targeted actions to be taken to identify and
36 manage potential sources of pathogen transmission in localised communities.

37 Keywords: COVID-19, near-to-source, normalisation, SARS-CoV-2, wastewater-based
38 epidemiology

39 **1. INTRODUCTION**

40 Wastewater-based epidemiology (WBE) is a promising tool for complementary surveillance
41 of infectious diseases and provision of early warning of disease outbreaks (Sims and
42 Kasprzyk-Hordern 2020), and has received considerable interest during the global COVID-19
43 pandemic (e.g. Bivins et al. 2020, Gonzalez et al. 2020, Polo et al. 2020). Since a significant
44 proportion of individuals infected with the severe acute respiratory syndrome coronavirus 2
45 (SARS-CoV-2) shed the virus' ribonucleic acid (RNA) in their faeces (Medema et al. 2020),
46 SARS-CoV-2 RNA concentrations in wastewater can be used to provide an indicator of the
47 disease prevalence without relying on clinical testing data. Wastewater networks can, thus, be

48 viewed as a reflection of the microbiome of the population using the upstream systems
49 (Newton et al. 2015). This is beneficial as wastewater surveillance is independent from
50 people participating in testing (a testing bias towards symptomatic individuals and testing
51 reluctance may result in underrepresentation of the infected in clinical test data).

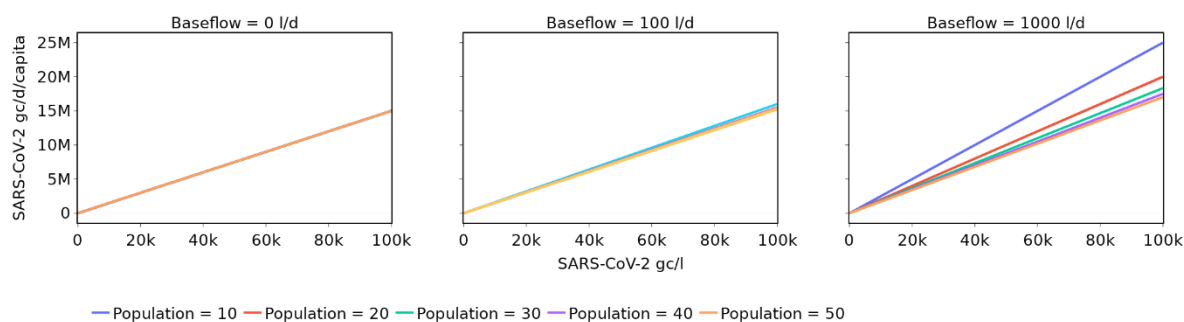
52 To date, SARS-CoV-2 related wastewater surveillance projects have been implemented in at
53 least 54 countries, covering more than 2,000 different sites (University of California 2021).
54 The European Union has recognised WBE as a tool to address emerging and future public
55 health issues (European Commission 2020), and Member States have been mandated to
56 engage with the development of the European Sewage Sentinel System for SARS-CoV-2,
57 which will provide systematic surveillance of SARS-CoV-2 and its variants in EU
58 wastewaters (Gawlik et al. 2021). In England, WBE for SARS-CoV-2 surveillance, led by the
59 Joint Biosecurity Centre and Defra Group under the Environmental Monitoring for Health
60 Protection (EMHP) programme, covers in excess of 500 sites as of June 2021.

61 Wastewater samples are commonly collected at sewage treatment works (STWs), thereby
62 providing an indicator of SARS-CoV-2 prevalence across the entire STW catchment – in the
63 Netherlands, for example, samples are taken daily at every STW in the country and analysed
64 for presence of SARS-CoV-2 (Dutch Water Sector, 2020). However, in-network or near-to-
65 source sampling can provide greater resolution and, potentially, additional insights: If the
66 population upstream of the sampling point is smaller, then better targeted actions can be
67 taken to address and mitigate any outbreak detected by monitoring of the wastewater.

68 Application of wastewater surveillance at a building scale, for example, can aid management
69 of outbreaks in discrete populations, and may be beneficial in high risk settings such as
70 schools, prisons and critical points in the food supply chain (Wade et al. 2020).

71 Near-to-source wastewater monitoring, employed at a hospital building, was able to detect a
72 single asymptomatic individual among as many as 400 residents (Karthikeyan et al. 2021).
73 The same study also found that trends in the number of cases could be successfully captured,
74 with a strong correlation between the wastewater SARS-CoV-2 concentration (gene copies
75 (gc)/l) and the number of active COVID-19 patients identified.

76 However, challenges remain in the interpretation and use of wastewater SARS-CoV-2 data as
77 an indicator of prevalence – one of which is the impact of variation in the upstream
78 population size. The importance of accounting for fluctuating population sizes in WBE has
79 previously been highlighted and investigated in the context of applications such as illicit drug
80 monitoring (Been et al. 2014). However, in wastewater SARS-CoV-2 monitoring, the value
81 considered as an indicator of prevalence is still typically reported as a concentration (gc/l)
82 (i.e. not normalised with respect to population like prevalence values) (e.g. Karthikeyan et al.
83 2021, Saththasivam et al. 2021, Prado et al. 2021), or even just a binary presence or absence
84 of SARS-CoV-2 RNA (e.g. Gibas et al. 2021). Using concentration as an indicator of
85 prevalence may be reasonable for monitoring efforts at a city-scale (i.e. those at a STW)
86 when the population size and wastewater dilution are relatively static, as a constant average
87 SARS-CoV-2 load per capita would then correspond to a constant SARS-CoV-2
88 concentration (assuming constant wastewater production per capita). It is not valid, however,
89 when population is highly variable and/or there are significant dilution events, since a
90 constant average SARS-CoV-2 load per capita would then yield a variable SARS-CoV-2
91 concentration. This is illustrated in Figure 1, which shows the theoretical relationship
92 between SARS-CoV-2 concentration, SARS-CoV-2 load per capita, population size and
93 baseflow (i.e. flow not attributed directly to the population), based on a mass balance
94 approach and assuming a per capita wastewater production of 150 l/d (further details in the
95 Supplementary Information).



96

97 *Figure 1. Relationship between SARS-CoV-2 concentration and daily load per capita for*
98 *different populations and base flows*

99 The impacts of variable population are likely to be exacerbated when interpreting the results
100 for near-to-source sites, since population movement will not be averaged out to the same
101 extent that it is at the STW level. This is especially important for near-to-source sites that are
102 subject to significant daily variability in upstream population (for example due to a lack of
103 activity at weekends or outside of academic term time at educational sites), and approaches
104 are required to ensure that the effects of variable populations are not overlooked.

105 The first aim of this study, therefore, is to investigate methods by which the effects of
106 variation in the upstream population at a highly dynamic near-to-source site can be accounted
107 for, and to build a better understanding of the impact of highly variable populations at a near-
108 to-source site on the insights into SARS-CoV-2 prevalence from WBE. Ammoniacal nitrogen
109 ($\text{NH}_3\text{-N}$) and orthophosphate (PO_4^{3-}) in the wastewater are considered as potential indicators
110 of population dynamics and for population normalisation. The work presented here is
111 specifically focused on a university campus case study, where footfall varies considerably
112 throughout the year (especially during periods of lockdown), and where knowledge of
113 COVID-19 prevalence provided by WBE has significant potential to target actions to identify
114 positive cases and reduce transmission between the university and the wider community.

115 Whilst near-to-source surveillance captures wastewater from a smaller and better-defined
116 population than monitoring at a STW level, the pool of potential candidates for the source of
117 any SARS-CoV-2 detected may still be large. The university monitored in this study, for
118 example, currently has approximately 20,900 students based at the campus and over 4,300
119 staff members. As such, the insights provided by WBE would be of greater assistance to
120 campus managers if the potential source(s) could be narrowed down further, e.g. to the most
121 probable building or buildings on site, so that better targeted public health actions may be
122 taken. The second aim of this study, therefore, is to investigate the potential of
123 complementary data to help direct response activities to the most appropriate locations within
124 the site from which wastewater is collected. This is achieved with the use of toilet flush data,
125 collated at washroom and building levels, which provide information on the major sources of
126 wastewater at the time of SARS-CoV-2 detection and the relative activity levels at different
127 locations on the campus.

128 Lastly, this study identifies samples collected during wet and dry weather periods to enable
129 exploration of the impact of wet weather periods on wastewater SARS-CoV-2 concentrations
130 and population normalisation. This aims to establish the value of wastewater monitoring in
131 wet versus dry periods, and address potential concerns arising due to the presence of heavily
132 diluted samples during wet weather.

133 It is intended that the results of this study will give improved confidence in the SARS-CoV-2
134 trends observed in wastewater from near-to-source sampling sites, and will enable better
135 targeted actions to identify and manage potential sources of pathogenic transmission in
136 localised communities.

137 2. METHODS

138 2.1. Case study site selection

139 The University of Exeter's Streatham Campus was monitored as a near-to-source pilot as part
140 of the UK National Surveillance Programme. Wastewater samples were collected in three
141 locations at the university: Two were downstream of student residential accommodation, and
142 one downstream of the main campus (consisting of multiple academic, administrative and
143 social buildings). This study focuses on the main campus monitoring site. The results are
144 expected to be of particular benefit here, since the number of people on campus is much more
145 variable than in accommodation blocks (due to working patterns and lecture scheduling, for
146 example), and population normalisation will thus have greater impact on the understanding of
147 prevalence. The maximum size of the potential upstream population is also considerably
148 larger, given that a large proportion of the university's 25,000 students and staff will have
149 access to the main campus buildings (although not at the same time), whereas only a small
150 proportion will be associated with individual accommodation sites. As such, any insights
151 provided by this study that enable the potential source of SARS-CoV-2 detected in the
152 wastewater to be narrowed down further will be much more valuable for the larger, main
153 campus site.

154 Furthermore, whilst the university does hold data on the number of students and staff that
155 have tested positive for SARS-CoV-2, there are no data on the total number of students and
156 staff using the campus each day (or whether those that have tested positive have actually been
157 on the main campus site on a given day) and, thus, prevalence cannot be calculated from
158 existing sources of information.

159 2.2. SARS-CoV-2 population normalisation

160 The daily wastewater SARS-CoV-2 load per capita (L_d , gc/capita/day) (i.e. a value that is
161 comparable with prevalence) can be calculated using Eq. 1, where C_d is the SARS-CoV-2

162 concentration (gc/l), Q_d is the daily flow rate (l/day), and N_d is the population size (in this
163 case, the number of individuals using the university campus).

$$L_d = \frac{C_d Q_d}{N_d} \quad \text{Eq. 1}$$

164 However, the population size on any given day is unknown and must, therefore, be estimated
165 if this equation is to be used.

166 There are multiple potential approaches to population estimation. In WBE, biomarkers
167 (substances excreted by humans) that have homogeneous excretion throughout a community
168 at low variance can be used as indicators of population size (Choi et al. 2018). These include,
169 for example, creatinine and coprostanol (Daughton 2012), and the cross-assembly phage
170 (crAssphage) (Wu et al. 2021). Water quality parameters, such as biochemical oxygen
171 demand (BOD), chemical oxygen demand (COD), total phosphorus (TP), total nitrogen (TN)
172 and ammonia or ammonium, may also be considered as population biomarkers (Xagorarakis
173 and O'Brien 2019, Nuijs et al. 2011): Been et al. (2014), for example, used measured
174 ammonium concentrations in conjunction with the expected daily per capita ammonium
175 discharge to estimate population size for the purposes of illicit drug monitoring, and Rico et
176 al. (2017) generated population estimates based on TN, TP, BOD and COD and typical daily
177 per capita discharges for each parameter.

178 Provided that daily flow rate (Q_d , l/day) and biomarker concentration (X_d , mg/l) data is
179 available for a historical period with known population (N_d), the daily discharge per capita of
180 a biomarker (x , mg/capita/day) can be estimated using Eq. 2, and the population on any given
181 day using Eq. 3.

$$x = \left(\frac{X_d Q_d}{N_d} \right) \quad \text{Eq. 2}$$

$$N_d = \frac{X_d Q_d}{x} \quad \text{Eq. 3}$$

182 Once the daily discharge per capita of a specific biomarker is known, concentrations of a
183 substance of interest (in this case, SARS-CoV-2 RNA) can then be population normalised
184 using Eq. 4, without calculating the population size and without any ongoing requirement for
185 flow rate data (based on substitution of Eq. 3 into Eq. 1):

$$L_d = \frac{C_d x}{X_d} \quad \text{Eq. 4}$$

186 Where L_d is the daily per capita load of SARS-CoV-2 at day d (gc/capita/day) and C_d is the
187 SARS-CoV-2 concentration in the wastewater at day d (gc/l).

188 Population estimates may also be generated using non-wastewater data: Thomas et al. (2017)
189 and Deville et al. (2014), for example, produced dynamic population estimates using mobile
190 phone communication data. With respect to the case study site in particular, there is ongoing
191 investigation into the potential use of wi-fi tracking for population monitoring under the
192 ‘*Riba to Reality*’ project (UKRI 2020); whilst there are no results available for use from this
193 yet, it is a promising development. Additional data available for the case study site that may
194 be indicative of population includes metered water and electricity supply and washroom-level
195 flush counts.

196 Each of these methodologies is subject to limitations, however, and there are some barriers to
197 implementation in a near-to-source study, where it should be noted that not everyone on site
198 will contribute to the wastewater collected since some visits may only be short. Whilst
199 mobile phone-based population estimates have the advantage that they can account for people
200 who spend only a short period of time in the monitored location, irrespective of whether they
201 produce any wastewater, this may be a disadvantage in near-to-source WBE as the population
202 estimate is likely to exceed the number of people that contribute to the wastewater being

203 sampled, and thus result in an underestimate of SARS-CoV-2 gene copies per capita. To
204 avoid this problem, a population normalisation methodology based on wastewater
205 characteristics is considered more appropriate for a near-to-source site with a highly dynamic
206 population.

207 For other non-wastewater indicators such as metered water, electricity supply and flush
208 counts, data is only available for a subset of buildings on campus. As there is insufficient
209 evidence that the occupancy dynamics of these buildings are representative of those of the
210 whole campus, it is inappropriate to use these for population normalisation (although
211 building-level population estimates may aid SARS-CoV-2 source identification, as discussed
212 in Section 2.3).

213 There are also potential disadvantages of using biomarkers in wastewater for population
214 normalisation. For the case study site, ammoniacal nitrogen and orthophosphate
215 concentrations are available, but it is recognised that nutrient concentrations will be affected
216 by industry (Xagorarakis and O'Brien, 2019) and, where there is significant industrial input
217 into the sewer network, this may lead to errors in population estimates (Nuijs et al. 2011).
218 Other studies, however, have concluded that use of such nutrients is appropriate (e.g. Choi et
219 al. 2018, Been et al. 2014), and Zheng et al. (2017) found population estimates based on
220 ammonia-nitrogen to show good agreement with estimates provided by wastewater treatment
221 plant operators with local knowledge. Furthermore, given that the focus of this study is a
222 near-to-source site, and it is known with a high degree of certainty that there are no industrial
223 inputs to the wastewater, these concerns are not considered a barrier in this study.

224 A remaining issue with this approach is the need to calculate the daily biomarker load
225 excreted per capita. Whilst previous studies have either estimated a site-specific value using
226 historical data (e.g. Rico et al. 2017) or used published values from the literature (e.g. Nuijs

227 et al. 2011), neither of these approaches is appropriate for this study. To calculate a site-
228 specific value, data is required from a period with known population; however, no
229 wastewater data was collected at the case study site prior to the pandemic (when the campus
230 could be assumed to be at full capacity, i.e. a known population size), and occupancy figures
231 during the period of wastewater monitoring are unknown. Use of daily values from literature
232 is also inappropriate since these are typically calculated at a STW level and thus capture
233 discharge from every occupant over a full 24 hours, whereas campus users are only present
234 for part of the day and their biomarker discharge will, therefore, be proportionately smaller.
235 Based on Eq. 4, and given that the daily biomarker discharge per capita (x) is constant at a
236 given site but cannot be calculated, this study calculates values that are proportional to the
237 SARS-CoV-2 load per capita (gene copies per mg of biomarker) instead of the actual load per
238 capita (gene copies per capita) (Eq. 5). These values are not comparable between sites;
239 however, they are sufficient for understanding the effects of population normalisation on
240 SARS-CoV-2 trends at a single site. Ammoniacal nitrogen and orthophosphate are both
241 considered as potential biomarkers.

$$L_t \propto \frac{C_d}{X_d} \quad \text{Eq. 5}$$

242 **2.3. Source identification**

243 As discussed in Section 2.2, supplementary data, including metered water supply, electricity
244 supply and flush counts, are available for a subset of buildings on campus, and may be used
245 to analyse occupancy dynamics within these buildings. Preliminary investigation shows very
246 poor agreement between building-level population estimates based on metered water and
247 electricity supply at the case study site (see Supplementary Information for example) –
248 potentially due to seasonal and other non-population-related influences – and, therefore, only
249 flush counts are considered further.

250 Given that flush count data relates directly to the production of wastewater, this information
251 can be used to assess the relative contribution of different buildings to the wastewater
252 sampled, and thus provide additional insights into the potential source(s) of any SARS-CoV-
253 2 RNA detected in the wastewater.

254 It has previously been shown that the number of flushes per occupant varies by building
255 (Melville-Shreeve et al. 2021); therefore, flush counts cannot be compared directly between
256 buildings and instead are used to estimate the occupancy of each building on any given day.
257 The maximum capacity of each building in which flush counts were monitored is shown in
258 Figure 2, and it is assumed that on weekdays during the autumn term of 2019 (pre-pandemic)
259 all buildings were operating at full capacity. The mean daily flush count per capita is then
260 estimated for each building using Eq. 6, and building occupancies during the wastewater
261 monitoring period are estimated using Eq. 7.

$$f_i = \frac{\sum_{d \in T} F_{i,d}}{|T|} \quad \text{Eq. 6}$$

$$P_{i,d} = \frac{F_{i,d}}{f_i} \quad \text{Eq. 7}$$

262 Where f_i is the mean daily per capita flush count for building i , $F_{i,d}$ is the total flush count for
263 building i on day d , T denotes the set of term time days where building occupancy is assumed
264 to equal building capacity, and $P_{i,d}$ is the estimated population of building i on day d .

265 This methodology is not used to estimate the total population of the campus on a daily basis
266 for the purposes of population normalisation as only a subset of the buildings have flush
267 counts monitored. As significant variability is found between buildings for the daily per
268 capita flush counts and the population estimates (see Section 3.3), it is not considered
269 appropriate to assume that the population dynamics for the monitored buildings are indicative
270 of those for the whole campus. Based on flush count data, it is not possible to either eliminate

271 or target any of the buildings that are not monitored as a potential source of any SARS-CoV-
272 2 detected in the wastewater; however, the population estimates provided can be used to
273 prioritise/deprioritise each of the buildings with flush count data available in the search for
274 the source of the SARS-CoV-2. A set of flush data for every toilet on campus would yield
275 additional value as a future extension to this study.

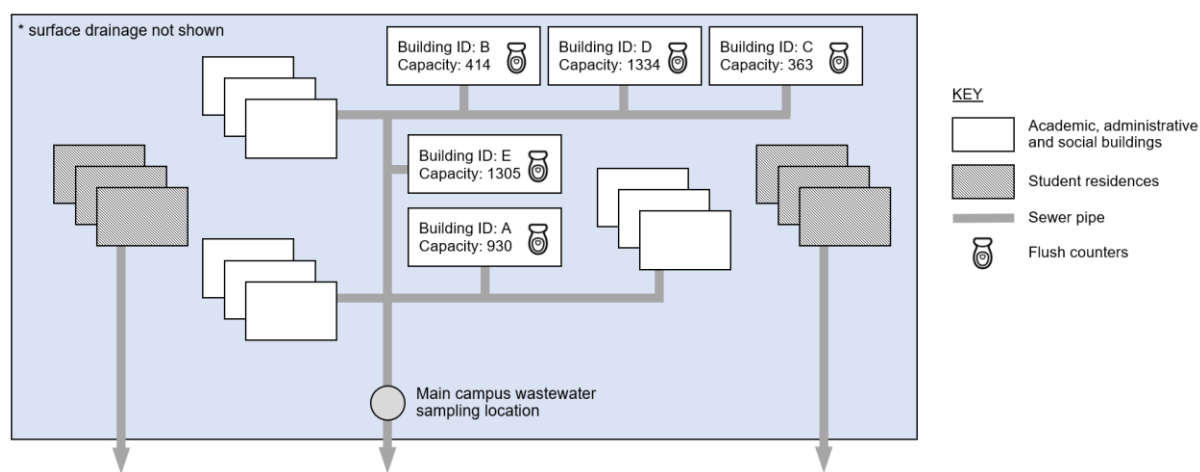
276 **2.4. Wet and dry weather days**

277 Days are classified as wet weather or dry weather (or neither) based on Environment Agency
278 rainfall data for the nearest rain gauge (station ID 45184, located approximately 4km from
279 the case study site) (Environment Agency, 2021). A day is considered a dry weather day if no
280 rainfall has been recorded on that day or in the previous six hours. A day is considered a wet
281 weather day if the total rainfall for that day exceeds 2mm (with the threshold set based on the
282 depth of rainfall expected to cause runoff from impervious surfaces (Ladson 2019)). Days
283 that fall into neither category (i.e. have rainfall but less than 2mm, or have had rainfall in the
284 preceding six hours) are not classified as either wet or dry.

285 **2.5. Data collection and handling**

286 This study incorporates both wastewater data and supplementary data relating to occupancy
287 of the campus, and a schematic of the case study site illustrating the data collection locations
288 is given in Figure 2. Wastewater samples were collected downstream of the academic,
289 administrative and social buildings (wastewater from student residences on campus are not
290 contributory), and they were analysed to provide concentrations of SARS-CoV-2,
291 ammoniacal nitrogen and orthophosphate. Wastewater flow rate was measured at the same
292 location. Flush counts were collected from five buildings, which have a combined occupancy
293 of 4,346 at full capacity.

294 Further detail on the data collection methodologies and data handling for each data type is
295 provided in the following sections.



296
297 *Figure 2. Case study site schematic, illustrating sampling locations and wastewater flow*
298 *from campus buildings. Surface drainage is omitted for clarity.*

299 **2.5.1. Wastewater characteristics**

300 Methodologies for the SARS-CoV-2, ammoniacal nitrogen and orthophosphate data
301 collection and processing are summarised below. Further details are available in Wade et al.
302 (2020) and Hoffman et al. (2021).

303 Composite wastewater samples were collected over a 24-hour period using HACH AS950
304 autosamplers (100ml every 15 minutes); these were then retrieved and kept at 4°C to prevent
305 degradation of RNA during transportation to a laboratory for analysis. Quantitative
306 Polymerase Chain Reaction with a reverse transcriptase step (RT-qPCR) was used to quantify
307 the N1 gene from the SARS-CoV-2 virus in the wastewater samples. This provides a
308 measurement of the number of RNA copies in the sample, which is reported as gene copies
309 per litre of wastewater sample collected. The practical limit of detection (LOD) is 133 gc/l.
310 No adjustments for analytical efficiency are applied.

311 Concentrations of ammoniacal nitrogen and orthophosphate were determined using
312 colorimetric assays. In both cases, outliers (more than three standard deviations from the
313 mean) are omitted from further analyses, since these may originate from a sample that was
314 poorly mixed and thus unrepresentative of the average wastewater composition at that time.
315 These represent 0% of the ammoniacal nitrogen measurements and 1% of the orthophosphate
316 measurements (two samples) in the analysis period.

317 The level and velocity of wastewater in the sewer were monitored at two-minute intervals in
318 the same location as collection of wastewater samples, and were used to calculate wastewater
319 flow rates. To enable comparison with the wastewater constituent concentrations measured in
320 the composite samples, flow data is resampled to provide daily values.

321 **2.5.2. Flush counts**

322 Flush monitoring systems were installed in washrooms at the University of Exeter, as
323 described by Melville-Shreeve et al. (2021a, 2021b). These captured flush data for
324 washrooms in six buildings, covering a total of 38 washrooms, and provided real time flush
325 counts for 119 toilets (approximately 18% of toilets on campus). Wheelchair-accessible
326 washrooms were omitted for operational reasons, but all other washrooms in the selected
327 buildings were monitored. The selected buildings represent typical university department
328 buildings. The monitoring system provided high resolution data, with either a ‘zero’ (no
329 flush) or a ‘one’ (flush) recorded every minute for every toilet.

330 In the case of any faults in the system, gaps may appear in the data and give the impression of
331 a reduced daily flush count if not identified and accounted for. Therefore, for quality
332 assurance purposes, daily data completeness (the total number of data points collected,
333 expressed as a percentage of the expected number of data points for the time period) is
334 evaluated at a washroom level. Since estimating flush counts for periods with insufficient

335 completeness would provide results with unknown accuracy, all washrooms with less than
336 90% data completeness during the wastewater monitoring period are omitted from further
337 analysis. This leaves 22 washrooms, covering five buildings, as illustrated in Figure 2.
338 Additionally, any day with less than 90% data completeness in any washroom is omitted
339 from further analysis. No outliers (more than three standard deviations from the mean) are
340 present in the total daily flush counts.

341 Detailed data completeness results for individual washrooms are provided in the
342 Supplementary Information (Figure S1), and days omitted from analyses due to insufficient
343 site-level flush count data availability are identified in Figure 3, Section 2.6.

344 Since flush counts are required to provide an indication of building occupancy, flushes
345 attributable to cleaning are identified and removed from the daily building level totals (as
346 these are not related to the building occupancy and do not contribute to SARs-CoV-2,
347 ammoniacal nitrogen or orthophosphate measured in the wastewater). A washroom is
348 assumed to have been cleaned if every toilet in the washroom has been flushed in a 15 minute
349 period at least once during the day; the number of flushes attributed to cleaning (a maximum
350 of one per toilet per day) is then subtracted from the total flush count. Further detail on the
351 impact of flush count adjustments for cleaning is provided in the Supplementary Information.

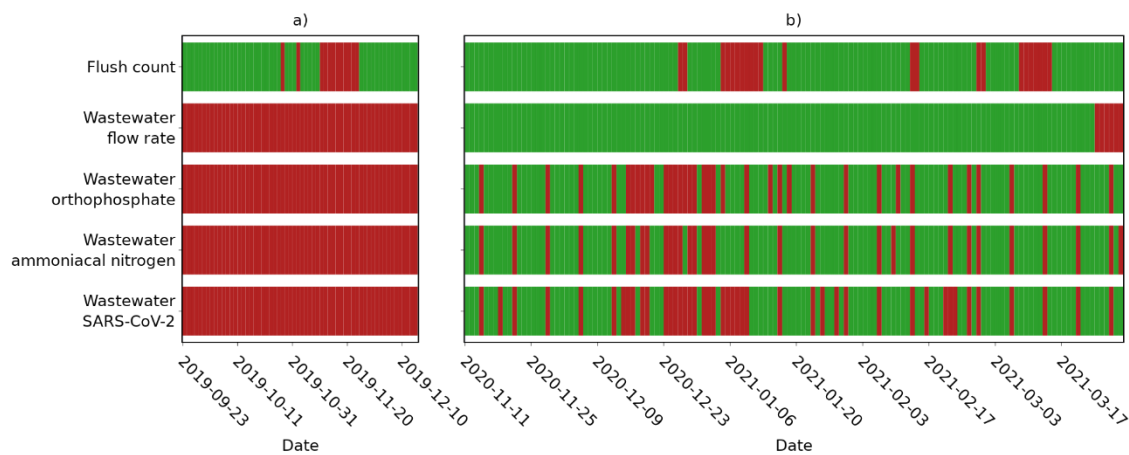
352 **2.6. Data availability and key dates**

353 Wastewater SARS-CoV-2, ammoniacal nitrogen and orthophosphate data is available from
354 11th November 2020, and sampling is ongoing. Calibrated wastewater flow rate data is
355 available from the same date, up until 23rd March 2021. Within these periods, wastewater
356 constituent data is typically available for six days a week, although sampling was reduced
357 during the university's Christmas break (12th December 2020 – 3rd January 2021), and flow
358 rate data is available for every day. Flush data was collected from 1st July 2019 until 29th

359 March 2021, although gaps are present where days have been omitted due to insufficient data
360 completeness.

361 The period 11th November 2020 to 29th March 2021, i.e. when both wastewater constituent
362 and flush data is available, is selected for analysis of the impact of population dynamics and
363 SARS-CoV-2 trends. Additional flush data from a period with assumed full occupancy
364 (weekdays during the 2019 autumn term) is also used in the source identification, as detailed
365 in Section 2.3.

366 Figure 3 shows the availability of each data source over these two periods, with data gaps
367 shown in red. Key dates within the wastewater monitoring period which may affect the
368 population on campus are summarised in Table 1.



369

370 *Figure 3. Daily data availability after cleaning, a) during a period of expected full*
371 *occupancy, and b) for the wastewater analysis period. Red indicates periods of missing data.*

372 *Table 1. Key dates in the wastewater SARS-CoV-2 analysis period*

Date	Description	Abbreviation
2 nd December 2020	End of second national lockdown	NL2-E
11 th December 2020	End of university term	T-E

24 th December 2020	Start of university closure period	CP-S
1 st January 2021	End of university closure period	CP-E
4 th January 2021	Start of university term	T-S
6 th January 2021	Start of third national lockdown	NL3-S

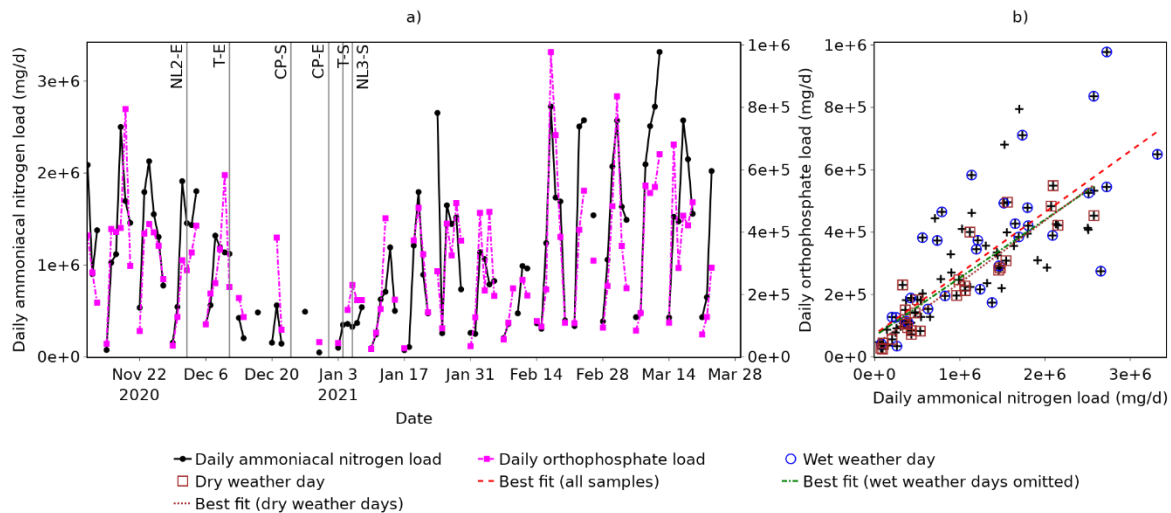
373

374 **3. RESULTS AND DISCUSSION**

375 **3.1. Population dynamics**

376 Population dynamics indicated by ammoniacal nitrogen and orthophosphate are analysed to
377 check that they appear plausible and to confirm that these biomarkers are a reasonable basis
378 for population normalisation for the case study site.

379 Site-specific values for the daily discharge per capita of ammoniacal nitrogen and
380 orthophosphate are unknown and cannot be calculated with the available information, and
381 thus absolute population cannot be estimated. The total biomarker load is, however,
382 proportional to population (based on Eq. 3) and can be used to illustrate population dynamics
383 and trends. Daily loads of ammoniacal nitrogen and orthophosphate, each of which is
384 expected to be proportional to the population on campus, are therefore shown in Figure 4.
385 These are calculated using the measured wastewater flow rate and biomarker concentration
386 ($Q_d X_d$).



404 The correlation between ammoniacal nitrogen and orthophosphate loads (Figure 4b) is strong
405 but not perfect (Pearson correlation coefficient $r = 0.792$), indicating that the population
406 estimates and population normalisation results yielded by each will differ. This may be
407 attributed to variations in the per capita discharge of each biomarker and the potential
408 presence of additional sources of either, or both, and is not unexpected given previous studies
409 have also reported discrepancies between population estimates based on different wastewater
410 constituents (e.g. Nuijs et al. 2011). Given that no population estimates from alternative
411 sources are available for reference, it is not possible to determine whether ammoniacal
412 nitrogen or orthophosphate provides a more accurate representation of population dynamics
413 and, therefore, both are considered in the following normalisation (Section 3.2).

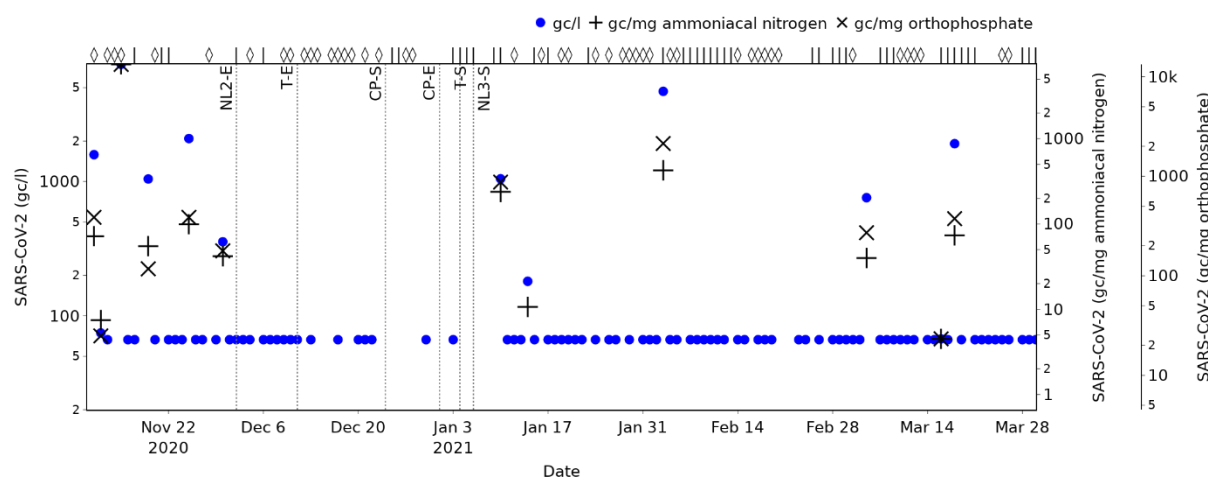
414 If samples taken during periods of wet weather are omitted, the correlation increases
415 marginally ($r = 0.804$), and if only samples taken on dry weather days are included, the
416 correlation increases further ($r = 0.917$). This suggests that confidence in population
417 estimates and population normalisation should be greatest during dry weather days; however,
418 population normalisation is not restricted to these days since this would prevent normalisation
419 of over 70% of samples, and there is still reasonable agreement between population estimates
420 based on the different biomarker loads even on wet weather days ($r = 0.758$).

421 Due to the topology and geology of the site (steep sandstone and mudstone), ingress to the
422 sewer (infiltration of groundwater) is not expected to be a major contributor to the wastewater
423 sampled; however, a similar study in an area with a high water table and a poorly maintained
424 pipe network may yield different results due to the increased impact of wet weather periods.

425 **3.2. Impact of population normalisation on wastewater SARS-CoV-2 trends**

426 SARS-CoV-2 RNA concentrations measured in wastewater samples at the case study site
427 over the period 11th November 2021 to 29th March 2021 are shown in Figure 5 (blue lines).

428 Samples with a concentration below the LOD are displayed with a value of 20 gc/l so that
 429 general trends can be observed; this value is half the LOD and selected based on the
 430 assumption that all values below the LOD are equally probable. The concentrations in Figure
 431 5 show that there is only intermittent detection of SARS-CoV-2 during the monitoring period
 432 (13% of samples collected contain a detectable level of SARS-CoV-2), with the most
 433 frequent detection and the maximum concentration both occurring in December 2020.
 434 All SARS-CoV-2 RNA concentrations above the LOD are normalised with respect to
 435 ammoniacal nitrogen and orthophosphate, and shown with '+' and 'x' symbols respectively
 436 in Figure 5, providing values that are expected to be directly proportional to the daily SARS-
 437 CoV-2 load per capita. To aid comparison of relative magnitudes with different units, y-axes
 438 are scaled so that the maximum gc/l, gc/mg ammoniacal nitrogen and gc/mg orthophosphate
 439 (all on 15th November 2020) all appear at the same level, and similar for the minimum
 440 detected values (all on 16th March 2021).



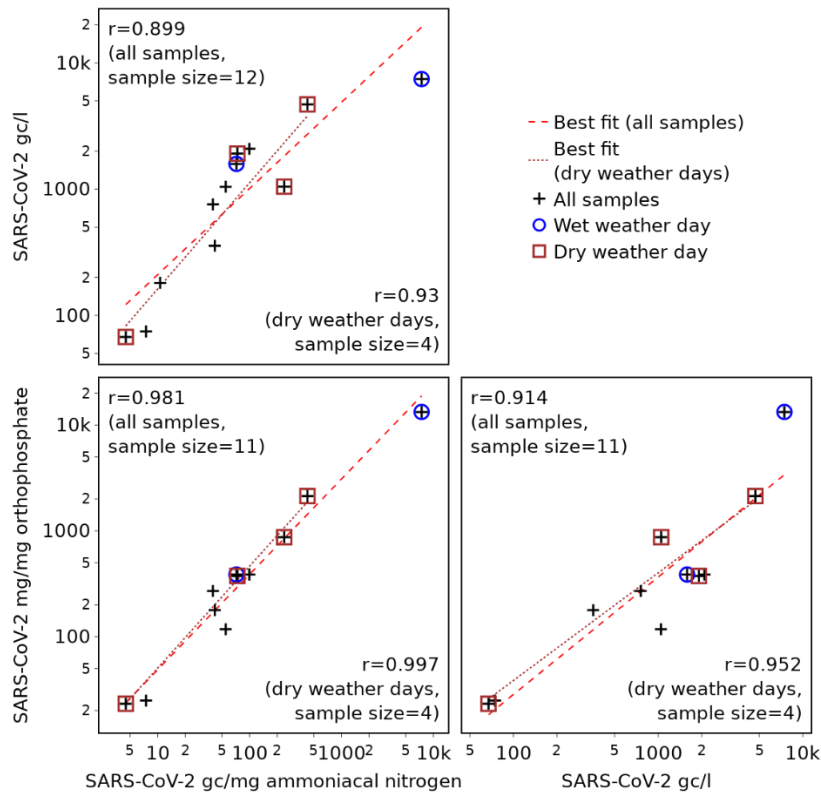
442 *Figure 5. Wastewater SARS-CoV-2 concentrations and normalised values over the*
 443 *wastewater monitoring period. Wet weather days are indicated with a '◇' symbol, and dry*
 444 *weather days with a '|' (top axis).*

445 Figure 5 shows that population normalisation (whether based on ammoniacal nitrogen or
446 orthophosphate) changes the picture of SARS-CoV-2 prevalence trends provided by the
447 wastewater monitoring. The wastewater samples taken on 19th November 2020 and 10th
448 January 2021, for example, have very similar SARS-CoV-2 RNA concentrations (1045 and
449 1048 gc/l respectively), but there is significant difference in their normalised values: Based
450 on these, the daily per capita load is shown to be 333 % (based on ammoniacal nitrogen for
451 normalisation) or 644% (based on orthophosphate for normalisation) higher on 10th January
452 than on 19th November. This compares with an increase of just 0.3% in concentration.
453 Neither of these samples were taken on a wet weather day, so the change cannot be attributed
454 to dilution effects. This suggests, therefore, that population normalisation can have significant
455 impact on the understanding of the relative severity of peaks in SARS-CoV-2, and that the
456 presence of similar SARS-CoV-2 RNA concentrations may give a false impression of similar
457 levels of prevalence.

458 This is supported by analysis of the rankings of the SARS-CoV-2 RNA concentrations and
459 population normalised values. Whilst the rankings based on gc/mg ammoniacal nitrogen and
460 gc/mg orthophosphate both have a strong correlation with the SARS-CoV-2 RNA
461 concentration (Spearman's rank correlation coefficients of 0.95 and 0.91 respectively), the
462 sample on 10th January is an example of where population normalisation significantly
463 changes the ranking (from 6th / 12 based on concentration, to 3rd based on gc/mg ammoniacal
464 nitrogen or gc/mg orthophosphate) and may alter the importance placed on the measurement
465 and any response actions that may be considered.

466 A comparison of concentrations and population normalised values for all samples with a
467 detectable level of SARS-CoV-2 is presented in Figure 6, illustrating more clearly the
468 relationship between each metric and which measurements correspond to dry and wet

469 weather days. In each case, linear lines of best fit and Pearson correlation coefficients shown
 470 are calculated based on the \log_{10} values.



471

472 *Figure 6. Correlations between wastewater SARS-CoV-2 concentrations and normalised*
 473 *values.*

474 Figure 6 shows that, although strongly correlated, the SARS-CoV-2 concentration (gc/l) is
 475 perfectly linearly related to neither the gc/mg ammoniacal nitrogen ($r = 0.899$) nor the gc/mg
 476 orthophosphate ($r = 0.914$). This reinforces the assertion that population normalisation alters
 477 the SARS-CoV-2 prevalence trends provided by WBE, since the relative magnitude (with
 478 respect to the full data set) of all SARS-CoV-2 concentrations and normalised loads would be
 479 unaltered only if there is a perfect, linear relationship ($r = 1$) between SARS-CoV-2
 480 concentration and the load per unit mass of biomarker. Whilst it may appear that the
 481 difference in the trends is likely to be minor, due to the strong correlation, the difference

482 between the linear best fit lines and the normalised values is in fact considerable (Root Mean
483 Square Percentage Error (RMPSE) = 83.7% for SARS-CoV-2 gc/mg ammoniacal nitrogen,
484 and RMSPE = 89.7% for SARS-CoV-2 gc/mg orthophosphate.

485 Comparison of the two population normalised values (gc/mg ammoniacal nitrogen and gc/mg
486 orthophosphate) shows that these exhibit a very strong correlation ($r = 0.981$), suggesting that
487 both will provide a similar understanding of trends in daily per capita loads of SARS-CoV-2.

488 Given that there is greater confidence in population normalisation using samples from dry
489 weather days (Section 3.1), correlation coefficients are also calculated based on dry weather
490 samples only. Again, these show there to be a very strong (nearly perfect) correlation
491 between the two population normalised metrics ($r = 0.997$), and a weaker correlation between
492 these and the SARS-CoV-2 RNA concentration ($r = 0.930$ and 0.952 for normalisation based
493 on ammoniacal nitrogen and orthophosphate respectively). These correlation coefficients are
494 based on only four samples (as this is the total number of samples with a SARS-CoV-2
495 concentration above the LOD collected on dry weather days) and, therefore, the confidence
496 intervals (detailed fully in the Supplementary Information, Table S1) are wider than for the
497 correlation coefficients based on all samples. However, the correlation coefficients still
498 support the conclusion that population normalisation alters the SARS-CoV-2 trends provided
499 by WBE, as there is not a perfect linear relationship between SARS-CoV-2 concentration and
500 SARS-CoV-2 normalised by either biomarker. They also support the suggestion that
501 normalisation using either metric is similarly beneficial, as the correlation between SARS-
502 CoV-2 gc/mg ammoniacal nitrogen and gc/mg orthophosphate is very strong even when
503 considering the full confidence interval ($0.867 \leq r \leq 1.000$ for a significance level of 0.05).

504 3.3. Potential sources of SARS-CoV-2

505 Estimated daily flush counts per occupant for each of the monitored buildings during the
506 period of assumed full occupancy, as calculated using Eq. 6 and required to estimate dynamic
507 building occupancies for SARS-CoV-2 source identification, are summarised in Table 2. The
508 total flush counts during the assumed full occupancy period, from which these values are
509 calculated, are provided in the Supplementary Information (Figure S2). The coefficients of
510 variation indicate that there will be a high degree of certainty in population estimates based
511 on the estimated flush count per occupant for Building C, and greatest uncertainty for
512 Building A. The significant variability between buildings in the mean daily flush count per
513 capita broadly matches the trends observed by Melville-Shreeve et al. (2021a) and may be
514 explained by variation in what each building is used for. Buildings that are used for teaching
515 (A, B, D and E), for example, may have a very high capacity but a relatively low occupancy
516 duration for each individual, and thus a low mean number of flushes per occupant;
517 conversely, each user of a building that is predominantly used for offices (C) may spend a
518 longer period of time in the building and therefore contribute more flushes. However, any
519 changes in building use as a result of the pandemic (such as a building with teaching space
520 being used only for research) may contribute further uncertainty in the daily per capita
521 flushes and associated occupancy estimates.

522 *Table 2. Daily per capita flush counts used to estimate dynamic population, calculated based*
523 *on building capacities and flush counts during a full occupancy period*

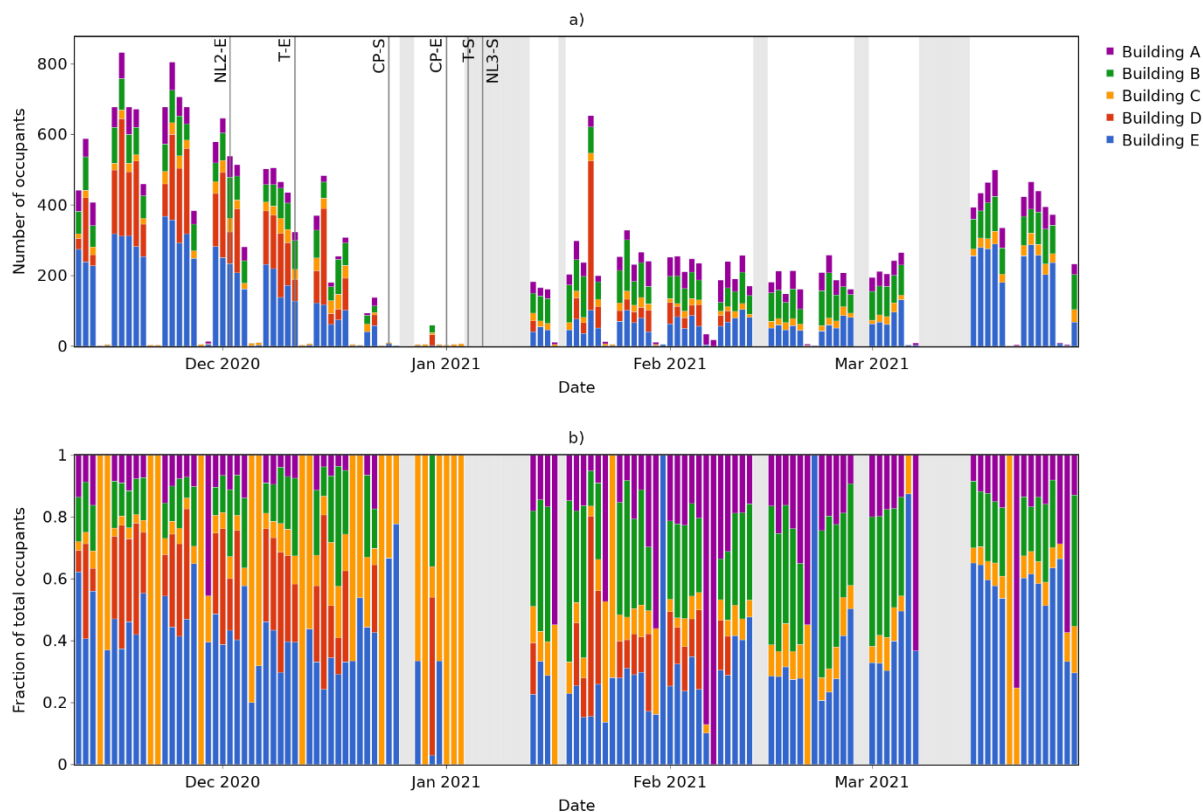
Building	Estimated number of flushes per occupant per day		
	Mean	Standard deviation	Coefficient of variation
A	0.336	0.101	0.30
B	0.283	0.061	0.21
C	2.033	0.172	0.08
D	0.033	0.006	0.19
E	0.578	0.113	0.20

524

525 Building-level occupancy estimates during the wastewater monitoring period, based on total
526 flush counts (available in the Supplementary Information, Figure S4) and the estimated
527 number of flushes per occupant per day (Table 2) are provided in Figure 7. The estimated
528 number of occupants in each building (Figure 7a) exhibits a clear weekly pattern, with near
529 zero occupancy in all buildings at weekends. Similarly to the campus-level population trends
530 indicated by the wastewater biomarker loads (Figure 4a), building occupancy drops
531 throughout December, and is lowest between the start of the Christmas closure period (CP-S)
532 and the following start of term (T-S).

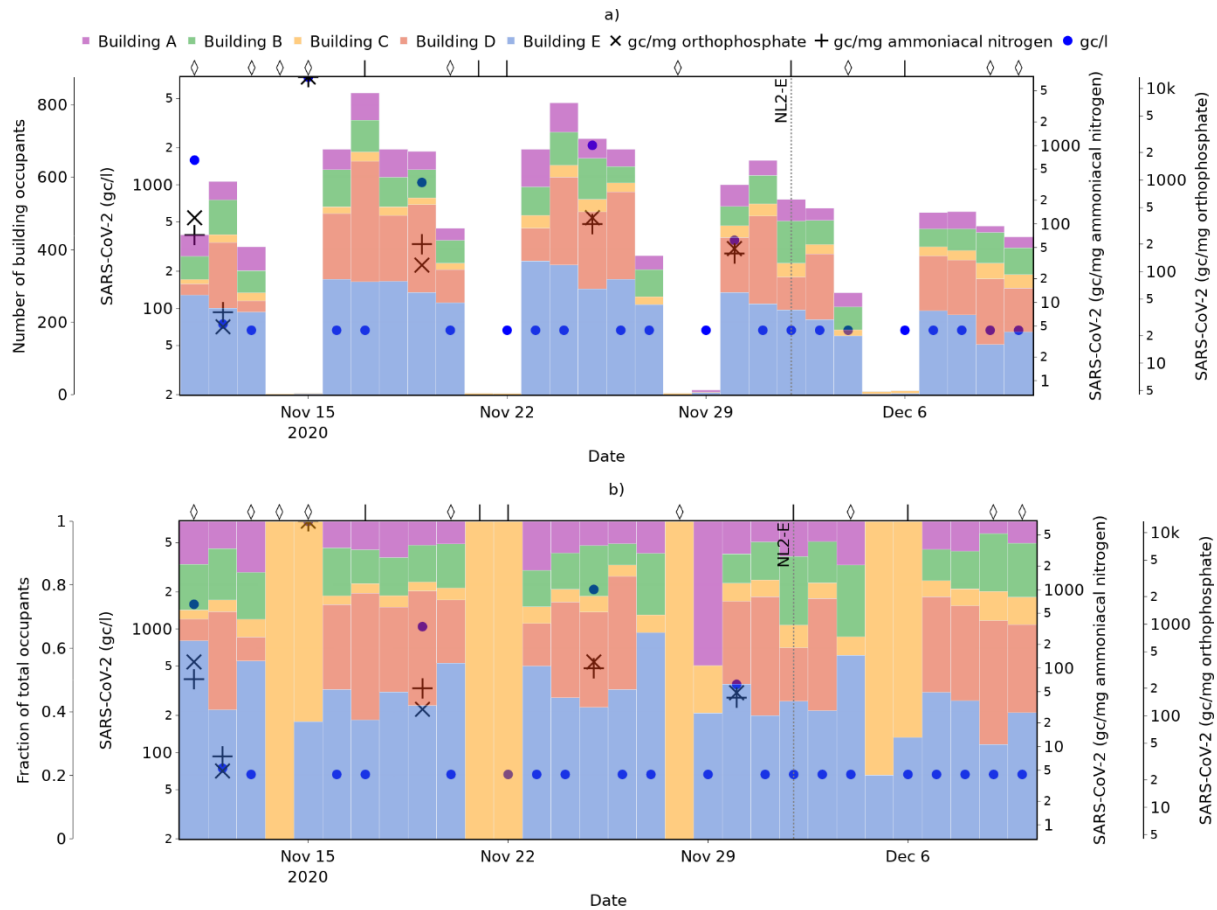
533 Figure 7b shows how occupants are distributed between the monitored buildings and,
534 therefore, the relative contribution of each building to the wastewater monitored for SARS-
535 CoV-2. This is particularly insightful on days with low total occupancy numbers, showing for
536 example that there are several days where occupants are detected only in Building C. Whilst
537 this does not guarantee that there was no-one in the other monitored buildings, it does mean
538 that nobody in them contributed to the wastewater being sampled, and thus these buildings

539 can be eliminated when searching for the source of any SARS-CoV-2 detected in the
540 wastewater on these days.



541
542 *Figure 7. Building occupancy estimates based on flush count data: a) Absolute occupancy;*
543 *and b) Relative occupancy (fraction of total occupants in monitored buildings).*

544 To illustrate the potential benefit of these building-level occupancy estimates for SARS-CoV-
545 2 source identification, Figure 8 shows the wastewater SARS-CoV-2 metrics for the one-
546 month period in which detection was most frequent (and, thus, source identification is of
547 greatest potential benefit) overlaid on corresponding occupancy estimates.



548

549 *Figure 8. Estimated occupancy of buildings with monitored flush counts and measures of*
 550 *site-level SARS-CoV-2 in wastewater during the period 11th November to 11th December*
 551 *2020, showing a) absolute occupancy; and b) relative occupancy. Wet weather days are*
 552 *indicated with a '◇' symbol, and dry weather days with a '|'.*

553 On 15th November 2020, the highest wastewater SARS-CoV-2 levels seen at the university
 554 were recorded. Figure 8b shows that buildings A, B, and D can be removed as candidates for
 555 the source, since their estimated occupancy on this day was zero. Whilst there is uncertainty
 556 in most occupancy estimates due to variation in the building-specific daily flushes per capita
 557 values, there is greater certainty of any occupant that may have been present not contributing
 558 to the wastewater on days where the estimated occupancy is zero (i.e. zero flushes). However,
 559 it is noted that flush data was not captured at wheelchair-accessible washrooms, and thus the

560 possibility of users of these buildings contributing to the wastewater cannot be absolutely
561 ruled out.

562 Furthermore, Figure 8a shows that the occupancy of all buildings with monitored flush counts
563 were very low on this day. If it is assumed that the occupancy of these buildings is broadly
564 representative of occupancy across the campus on this day, then this suggests that the actual
565 number of people infected may also be very low, despite both the SARS-CoV-2
566 concentration and the population normalised values being very high (since the smaller the
567 population, the higher the per capita value resulting from a given SARS-CoV-2 load).

568 On days where specific buildings cannot be eliminated based on a zero-occupancy estimate,
569 building-level occupancy information may still aid efforts to trace the source of SARS-CoV-2
570 detected by enabling identification of buildings with the greatest occupancy and greatest
571 wastewater contribution. Figure 8b shows that on 11th November, for example, over 60% of
572 the total occupants of five monitored buildings were in Building E – hence, there is greater
573 probability of locating the infected individual(s) in this building, and this should be a higher
574 priority for targeted testing if capacity is limited.

575 **3.4. Future opportunities**

576 Multiple opportunities are identified to add value to outputs set out in this study. Specifically,
577 the installation of flush monitoring technology across all washrooms upstream of the
578 autosampler could enable a more definitive set of conclusions to be drawn. Disaggregation of
579 flush counts from male and female washrooms may also improve accuracy when assessing
580 the relative contribution of occupants in different buildings to the wastewater sampled, due to
581 the presence of (unmonitored) urinals in the male washrooms.

582 In addition, from an operational perspective, additional autosamplers could be installed (but
583 remain largely offline) at the outlet from each building. These could be sampled the day after

584 a positive signal is observed in at the main campus, enabling a single building to be pin-
585 pointed. Such measures could in turn enable patchwork closure of buildings when prevalence
586 exceeds a pre-defined threshold, minimising disruption in future waves of a pandemic.

587 **4. CONCLUSIONS**

588 This study has investigated the use of ammoniacal nitrogen and orthophosphate for
589 normalisation of SARS-CoV-2 detected in wastewater, to account for the impact of highly
590 variable populations at a near-to-source monitoring site; evaluated the impact of population
591 normalisation on the understanding of SARS-CoV-2 prevalence trends provided by the
592 wastewater data; demonstrated how complementary (non-wastewater) data sources can help
593 to inform a better targeted response; and explored the potential impact of wet weather periods
594 on the results. Key findings include:

- 595 • Population normalisation alters the trends in SARS-CoV-2 prevalence indicated by
596 WBE and, in a near-to-source site with a highly variable population such as a
597 university campus, it can reveal significant differences in prevalence between days
598 where recorded SARS-CoV-2 concentrations are very similar. Population
599 normalisation, therefore, is considered critical for providing a comprehensive
600 understanding of the results from WBE when population size is highly variable.
- 601 • Normalisation using either ammoniacal nitrogen or orthophosphate is similarly
602 beneficial, with both providing a similar (but not identical) understanding of
603 population dynamics and trends in population normalised SARS-CoV-2 in the
604 wastewater. This indicates that multiple biomarkers that are of questionable reliability
605 for population normalisation at a STW level due to their presence in industrial
606 discharges can be appropriate for near-to-source studies.

- 607 • Agreement between population estimates based on different biomarkers is greatest
608 when wet weather days are omitted, indicating that confidence in the results of
609 population normalisation should be greatest when the weather is dry. However, as
610 there is still a reasonable level of agreement on wet days, these can still provide
611 valuable information.
- 612 • Use of flush count data to estimate the occupancy of different buildings in within the
613 near-to-source site can enable priority locations for targeted testing to be identified
614 when SARS-CoV-2 is detected in the wastewater. This is particularly beneficial on
615 low occupancy days when no flushes are recorded in some buildings, so it is known
616 with certainty that no occupants of these buildings contributed to the wastewater in
617 which SARS-CoV-2 was detected.
- 618 • Technically feasible strategies to further advance this study have been set out. Such
619 solutions focus on yet more granular data acquisition including a wider deployment of
620 flush monitoring and short term autosampling being added at a building-level when
621 the main campus data suggests increased prevalence.

622 Lastly, it is noted that there were restrictions in place on mobility and/or student activities,
623 along with guidance to ‘work from home where possible’, for the majority this study period,
624 due to the COVID-19 pandemic. As such, the number of people using the main campus site
625 (and the difference between high and low occupancy periods) was considerably lower than
626 usual. For near-to-source sites with higher variability in population – and for the case study
627 site as restrictions are lifted and the number of people using the campus increases – the
628 importance of population normalisation is expected to be even greater.

629 **ACKNOWLEDGEMENTS**

630 This paper uses Environment Agency rainfall data licensed under the Open Government
631 Licence v3.0. The sampling, testing and data analysis of wastewater in England is funded by
632 the United Kingdom Government (Department of Health and Social Care).

633 **DISCLAIMER**

634 The views expressed in this paper are those of the authors and do not necessarily reflect the
635 views or policies of the Department of Health and Social Care.

636 **REFERENCES**

637 Been, F., Rossi, L., Ort, C., Rudaz, S., Delemont, O. and Esseiva, P. 2014. Population
638 Normalization with Ammonium in Wastewater-Based Epidemiology: Application to Illicit
639 Drug Monitoring. *Environmental Science & Technology* 48(14), 8162-8169.

640 Bivins, A., North, D., Ahmad, A., Ahmed, W., Alm, E., Been, F., Bhattacharya, P., Bijlsma,
641 L., Boehm, A., Brown, J., Buttiglieri, G., Calabro, V., Carducci, A., Castiglioni, S., Gurol, Z.,
642 Chakraborty, S., Costa, F., Curcio, S., de los Reyes, F., Vela, J., Farkas, K., Fernandez-Casi,
643 X., Gerba, G., Gerrity, D., Girones, R., Gonzalez, R., Haramoto, E., Harris, A., Holden, P.,
644 Islam, M., Jones, D., Kasprzyk-Hordern, B., Kitajima, M., Kotlarz, N., Kumar, M., Kuroda,
645 K., Rosa, G., Malpei, F., Mautus, M., McLellan, S., Medema, G., Meschke, J., Mueller, J.,
646 Newton, R., Nilsson, D., Noble, R., van Nuijs, A., Peccia, J., Perkins, T., Pickering, A., Rose,
647 J., Sanchez, G., Smith, A., Stadler, L., Stauber, C., Thomas, K., van der Voorn, T.,
648 Wigginton, K., Zhu, K., and Bibby, K. 2020. Wastewater-Based Epidemiology: Global
649 Collaborative to Maximize Contributions in the Fight Against COVID-19. *Environmental*
650 *Science & Technology* 54 (13), 7754-7757. Choi, P.M., Tschärke, B.J., Donner, E., O'Brien,
651 J.W., Grant, S.C., Kaserzon, S.L., Mackie, R., O'Malley, E., Crosbie, N.D., Thomas, K.V.

652 and Mueller, J.F. 2018. Wastewater-based epidemiology biomarkers: Past, present and
653 future. *Trac-Trends in Analytical Chemistry* 105, 453-469.

654 European Commission. 2020 Building a European Health Union: Reinforcing the EU's
655 resilience for cross-border health threats, –Brussels, 11.11.2020 COM(2020) 724.

656 Daughton, C.G. 2012. Real-time estimation of small-area populations with human
657 biomarkers in sewage. *Science of the Total Environment* 414, 6-21.

658 Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E., Blondel, V.D.
659 and Tatem, A.J. 2014. Dynamic population mapping using mobile phone data. *Proceedings*
660 *of the National Academy of Sciences of the United States of America* 111(45), 15888-15893.

661 Dutch Water Sector 2020 Nationwide Covid-19 sewer surveillance at all Dutch wwtps.
662 Available at: [https://www.dutchwatersector.com/news/nationwide-covid-19-sewer-](https://www.dutchwatersector.com/news/nationwide-covid-19-sewer-surveillance-at-all-dutch-wwtps)
663 [surveillance-at-all-dutch-wwtps](https://www.dutchwatersector.com/news/nationwide-covid-19-sewer-surveillance-at-all-dutch-wwtps). Accessed 24 June 2021.

664 Environment Agency 2021 Real time flood monitoring data: Archives. Available at:
665 <https://environment.data.gov.uk/flood-monitoring/archive>. Accessed 24 June 2021.

666 Gawlik, B., Tavazzi, S., Mariani, G., Skejo, H., Sponar, M., Higgins, T., Medema, G. and
667 Wintgens, T. (2021) SARS-CoV-2 Surveillance employing Sewage - Towards a Sentinel
668 System, EUR 30684 EN, Publications Office of the European Union, Luxembourg.

669 Gibas, C., Lambirth, K., Mittal, N., Juel, M.A.I., Barua, V.B., Brazell, L.R., Hinton, K.,
670 Lontai, J., Stark, N., Young, I., Quach, C., Russ, M., Kauer, J., Nicolosi, B., Chen, D.,
671 Akella, S., Tang, W.W., Schlueter, J. and Munir, M. 2021. Implementing building-level
672 SARS-CoV-2 wastewater surveillance on a university campus. *Science of the Total*
673 *Environment* 782.

674 Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M., Yetka, K., Thompson H., Keeling,
675 D., Mitchell, J., Gonzalez, D. 2020. COVID-19 surveillance in Southeastern Virginia using
676 wastewater-based epidemiology. *Water Research* 186, 116296.

677 Hoffmann, T., McIntyre-Nolan, S., Bunce, J., Grimsley, J., Hart, A., Lo Jacamo, A., Morvan,
678 M., Robins, K., Wade, M., Watts, G., Engeli, A. and Henderson, G. 2021 Current
679 environmental monitoring cannot constrain the effect of vaccines on SARS-CoV-2
680 transmission: Report for SAGE 08/04/2021.

681 Karthikeyan, S., Ronquillo, N., Belda-Ferre, P., Alvarado, D., Javidi, T., Longhurst, C.A. and
682 Knight, R. 2021. High-Throughput Wastewater SARS-CoV-2 Detection Enables
683 Forecasting of Community Infection Dynamics in San Diego County. *Msystems* 6(2).

684 Ladson, A. (2019) Approaches to Water Sensitive Urban Design. Sharma, A., Gardner, T.
685 and Begbie, D. (eds), pp. 209-228.

686 Medema, G., Heijnen, L., Elsinga, G., Italiaander, R. and Brouwer, A. 2020. Presence of
687 SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence
688 in the Early Stage of the Epidemic in The Netherlands. *Environmental Science & Technology*
689 *Letters* 7(7), 511-516.

690 Melville-Shreeve, P., Cotterill, S. and Butler, D. 2021a. Capturing high-resolution water
691 demand data in commercial buildings. *Journal of Hydroinformatics* 23(3), 402-416.

692 Melville-Shreeve, P., Cotterill, S., Newman, A. and Butler, D. 2021b. Campus Study of the
693 Impact of Ultra-Low Flush Toilets on Sewerage Networks and Water Usage. *Water* 13(4).

694 Newton, R., McLellan, S., Dila, D., Vineis, J., Morrison, H., Eren, A. and Sogin, M. 2015.
695 Sewage Reflects the Microbiomes of Human Populations. *mBio* 6(2), e02574-14.

696 Polo, D., Quintela-Baluja, M., Corbishley, A., Jones, D., Singer, A., Graham, D., Romalde, J.
697 2020. Making waves: Wastewater-based epidemiology for COVID-19 – approaches and
698 challenges for surveillance and prediction. *Water Research* 186, 116404.

699 Prado, T., Fumian, T.M., Mannarino, C.F., Resende, P.C., Motta, F.C., Eppinghaus, A.L.F.,
700 do Vale, V.H.C., Braz, R.M.S., de Andrade, J.D.R., Maranhao, A.G. and Miagostovich, M.P.
701 2021. Wastewater-based epidemiology as a useful tool to track SARS-CoV-2 and support
702 public health policies at municipal level in Brazil. *Water Research* 191.

703 Rico, M., Andres-Costa, M.J. and Pico, Y. 2017. Estimating population size in wastewater-
704 based epidemiology. Valencia metropolitan area as a case study. *Journal of Hazardous*
705 *Materials* 323, 156-165.

706 Saththasivam, J., El-Malah, S.S., Gomez, T.A., Jabbar, K.A., Remanan, R., Krishnankutty,
707 A.K., Ogunbiyi, O., Rasool, K., Ashhab, S., Rashkeev, S., Bensaad, M., Ahmed, A.A.,
708 Mohamoud, Y.A., Malek, J.A., Abu Raddad, L.J., Jeremijenko, A., Abu Halaweh, H.A.,
709 Lawler, J. and Mahmoud, K.A. 2021. COVID-19 (SARS-CoV-2) outbreak monitoring using
710 wastewater-based epidemiology in Qatar. *Science of the Total Environment* 774.

711 Sims, N. and Kasprzyk-Hordern, B. 2020. Future perspectives of wastewater-based
712 epidemiology: Monitoring infectious disease spread and resistance to the community level.
713 *Environment International* 139.

714 Thomas, K.V., Amador, A., Baz-Lomba, J.A. and Reid, M. 2017. Use of Mobile Device
715 Data To Better Estimate Dynamic Population Size for Wastewater-Based Epidemiology.
716 *Environmental Science & Technology* 51(19), 11363-11370.

717 UKRI 2020 From RIBA to Reality – Deep Digital Twin to enable Human-Centric Buildings
718 for a Carbon Neutral Future. Available at: <https://gtr.ukri.org/projects?ref=105874>. Accessed
719 24 June 2021.

720 University of California 2021 COVID19Poops Dashboard. Available at:
721 <https://www.covid19wbec.org/covidpoops19>. Accessed 24 June 2021.

722 van Nuijs, A.L.N., Mougél, J.F., Tarcomnicu, I., Bervoets, L., Blust, R., Jorens, P.G., Neels,
723 H. and Covaci, A. 2011. Sewage epidemiology - A real-time approach to estimate the
724 consumption of illicit drugs in Brussels, Belgium. *Environment International* 37(3), 612-621.

725 Wade, M., Jones, D., Singer, A., Hart, A., Corbishley, A., Spence, C., Morvan, M., Zhang,
726 C., Pollock, M., Hoffmann, T., Singleton, P., Grimsley, J., Bunce, J., Engeli, A. and
727 Henderson, G. 2020 Wastewater COVID-19 Monitoring in the UK: Summary for SAGE –
728 19/11/20.

729 Wu, J., Wang, Z., Lin, Y., Chen, J., Li, P., Liu, W., Wang, Y., Yao, C. and Yang, K. 2021.
730 Technical framework for wastewater-based epidemiology of SARS-CoV-2. *Science of the*
731 *Total Environment* 791, 148271.

732 Xagorarakis, I. and O'Brien, E. (2020) *Women in Water Quality*. *Women in Engineering and*
733 *Science*. D., O.B. (ed), Springer, Cham.

734 Zheng, Q.D., Lin, J.G., Pei, W., Guo, M.X., Wang, Z. and Wang, D.G. 2017. Estimating
735 nicotine consumption in eight cities using sewage epidemiology based on ammonia nitrogen
736 equivalent population. *Science of the Total Environment* 590, 226-232.

737