This accepted manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

The published version is available at:
Schmitt, R. J. P., Morgenroth, E., & Larsen, T. A. (2017). Robust planning of sanitation services in urban informal settlements: an analytical framework. Water Research, 110, 297-312. http://doi.org/10.1016/j.watres.2016.12.007

ACCEPTED MANUSCRIPT

Robust planning of sanitation services in urban informal settlements: an analytical framework.

Rafael Schmitt[1,2], Eberhard Morgenroth[2,3], Tove A. Larsen[2, *]

[1] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy.

[2] Eawag: Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

[3] ETH Zürich, Institute of Environmental Engineering, 8093 Zürich, Switzerland

* Corresponding author, Email: tove.larsen@eawag.ch, Phone: +41 58 765 5039.

ACCEPTED MANUSCRIPT

Abstract

New types of sanitation services are emerging to tackle the sanitation crisis in informal settlements. These services link toilet facilities to semi-decentralized treatment plants via frequent, road-based transport of excreta. However, information for the planning of such sanitation services is scarce, and their future operating conditions are highly uncertain. The key questions of this paper are therefore: a) what are the drivers behind success or failure of a service-based sanitation system in informal settlements and b) on what scales and under which conditions can such a system operate successfully? To answer these questions, already at an early stage of the planning process, we introduce a stochastic model to analyze a wide range of system designs under varying technical designs, socio-economic factors, and spatial condition. Based on these initial results, we design a sanitation service and use the numeric model to study its reliability and costs over a wide range of scales, i.e., system capacities, from very few to many hundred users per semi-decentralized treatment unit. Key findings are that such a system can only operate within a narrow, but realistic range of conditions. Key requirements are toilet facilities, which can be serviced rapidly, and a flexible workforce. A high density of facilities will also lower the costs. Under these premises, we develop a road-based sanitation service and model its functionality in different settings and under many scenarios. Results show that the developed sanitation system using a single vehicle is scalable (100 – 700 users), can provide reliable service, and can be cheap (< 1.5 c/p/day). Hence, this paper demonstrates opportunities for road-based sanitation in informal settlements and presents a quantitative framework for designing such systems.

1. Introduction

Deficient sanitation poses a major risk to human health (Prüss et al., 2002) and environmental sustainability (UNDP, 2014). The problem is pronounced in informal urban settlements where poor accessibility, the uncertain legal status of the inhabitants, and fast, unplanned development impede the implementation of sustainable sanitation systems. This is the case for the toilet infrastructure, and even more for the reliable transport of human waste, a precondition for later safe discharge and treatment (Katukiza et al., 2012; Lüthi et al., 2010).

ACCEPTED MANUSCRIPT

Traditionally, there are two approaches to sanitation system design: *On-site* treatment systems with long-term storage, or *off-site* treatment systems based on immediate transport of human waste via sewer lines (Katukiza et al., 2012), and subsequent (semi-) centralized treatment. Currently, *on-site* solutions serve up to 2.4 billion users, especially in low-income contexts where the installation of capital-intensive off-site solutions such as sewers is out of reach (WHO/UNICEF, 2012). Most *on-site* solutions require long retention of human waste and large storage volumes to ensure sufficient waste stabilization. This is a major challenge in informal settlements due to scarce space, unclear ownership, and the prohibitive costs of constructing appropriate on-site storage volumes (Paterson et al., 2007). A new generation of sanitation services has emerged to address the needs of the growing population that is served neither by *off-site* solutions (i.e. no access to sewers) nor *on-site* treatment (i.e. there is insufficient space for long-term storage). Sanitation services connect toilet facilities to semi-centralized treatment units through a frequent (e.g. weekly) demand-driven transport and service system (e.g. Loowatt, 2014; XRunner, 2014; Sanergy, 2015). The small amount of excreta accumulating between services allows the use of sealable storage containers for the safe and nuisance-free handling of excreta. Smaller amounts of excreta may be transported with standard vehicles on the existing road infrastructure, thus reducing investments while increasing system flexibility.

However, the required regular emptying of sanitation facilities and the transport of excreta to a treatment facility via the road network poses new challenges. Service systems are now common in informal settlements, e.g. for the distribution of consumer goods (Gates, 2010), but the level of service available for individual households is often limited (Kariuki and Jordan Schwartz, 2005), as is the information about these systems (UNHABITAT, 2013; Sharholy et al., 2008; Langenhoven and Dyssel, 2007). The planning of transport-based sanitation services is challenging with regard to: (1) the *a-priori* specification of system design parameters based on very sparse information (e.g. performance of the service vehicle, working habits); (2) expected large fluctuations in future operating conditions, e.g. fluctuating use of toilets or the future expansion of the sanitation system. All these unknown factors lead to a situation of deep uncertainty where neither potential future risks (e.g. system overloads or financial failure) nor the probability of their occurrence can be readily estimated. Such a situation of deep uncertainty impedes the application of typical planning approaches (Lempert et al., 2003). As a

ACCEPTED MANUSCRIPT

consequence, sanitation services in informal settlements are commonly based on *ad-hoc* organization, resulting in low service quality and ultimately in system failure (Murungi and van Dijk, 2014).

New approaches, such as bottom-up or robust decision-making (e.g. Brown et al., 2012), enable planning under conditions of deep uncertainty, e.g. regarding future climatic conditions. These approaches are increasingly promoted as an appropriate response to the planning challenges in emerging economies and dynamically growing mega-cities (Ranger and Garbett-Shiels, 2012). Bottom-up planning strategies are often based on numerical system models which allow the *a-priori* identification of risks and validation of systems functioning over a wide range of uncertain future operating conditions (Brown et al., 2012; Hallegatte, 2009). Such approaches identify trade-offs between competing goals as well as risks within and outside the system in order to develop robust designs and to foresee and avoid critical future conditions (Lempert et al., 2006; Lempert and Collins, 2007).

In this paper, we demonstrate how such a bottom-up approach can be beneficial for planning robust sanitation services in urban informal settlements. We implement a planning framework for sanitation services based on a flexible, stochastic model. The model is contextualized to specific informal settlements via a probabilistic analysis of spatial data derived from widely available, free-of-charge, satellite imagery. In the planning framework, the initial design of a sanitation service system is based on an exhaustive analysis of possible future operating conditions and potential risks, resulting in a system design that balances system cost and performance and is robust under a wide range of operating conditions. Performance of that system is then modelled under a large number of different conditions and for different system scales. The purpose of the modeling method is thus to identify a transport system, which will work under a large number of possible (uncertain) conditions.

The stochastic model allows the high uncertainty relating to sanitation planning at early planning stages to be addressed. Potential future critical conditions are identified *a-priori* by analyzing a high number of future scenarios. In the early steps of the modeling process, we only learn which parameters are most important and how they roughly influence the viability of a given transport system (e.g. we learn how costs depend on the emptying time of toilet facilities). In the later steps of the model, we include more site-specific information, e.g. about the transport distances in a given settlement.

The methodology was developed within a project on the Blue Diversion Toilet (Larsen et al., 2015), financed by the Bill & Melinda Gates Foundation. We thus use the financial constraints formulated in this call (maximum costs of 0.05 US$/person/day for the entire sanitation service) for measuring the possible success of a transport system. However, the methodology is equally valid for all other sanitation services intended for informal urban settlements, albeit the specific model may look different.

ACCEPTED MANUSCRIPT

2. Materials and Methods

The proposed framework is based on a stochastic numerical model of the day-to-day functioning of a sanitation service system (see Appendix 2). The modeling takes place in four steps starting with a rough screening procedure and ending with the modeling of specific scenarios in specific settings (Figure 1). In step 1, a numerical model for the sanitation system is used to analyze the functioning of a sanitation system for a wide range of technical, socio-economic, and spatial parameters. Parameter ranges are derived from literature or expert-based without considering a specific design. Performance targets are defined based on certain project premises (e.g., daily cost per user as compared to purchasing power). A sensitivity analysis identifies a) parameters that affect the system performance most and b) ranges of the identified sensitive parameters for which performance targets are matched (section 2.3.1). Step 2 (section 2.3.2) focusses the system design on these most sensitive parameters, and results in a preliminary design of the system. Step 2 also includes scenario development, i.e., possible future operating conditions (e.g., in terms of user numbers), design alternatives (e.g., in terms of vehicle selection), or operational strategies (e.g., in terms of employees and payment schemes). Step 3 (section 2.3.3) describes how spatial parameters for a specific informal settlement can be measured in a probabilistic manner from satellite imagery. These data are used to contextualize the model for a specific informal settlement. The probabilistic approach allows to contextualize the model even without knowledge on the actual system layout in that settlement (i.e., where toilet and treatment facilities are to be located), and it allows to consider scenario-specific decisions (from step 2), i.e., how many users to connect, or how often toilets need to be emptied. Step 4 (section 2.3.4) takes up the designs and scenarios from step 2, and uses the sanitation service model (see step 1) to model the functioning of the sanitation service for each scenario and in each specific setting defined in Step 3 over the entire life-time of the system. Step 4 estimates the system performance (e.g. in terms of costs and service capacity) and the probability of system failure for each scenario (failure or success is defined by performance targets set in step 1). Identifying scenarios under which the system fails or succeeds allows estimating the functioning of the system under a wide range of conditions and identifying conditions (e.g., in terms of system capacity, technical designs, or transport solutions) under which the system will be likely to operate successfully.

ACCEPTED MANUSCRIPT

NUSCRIPT

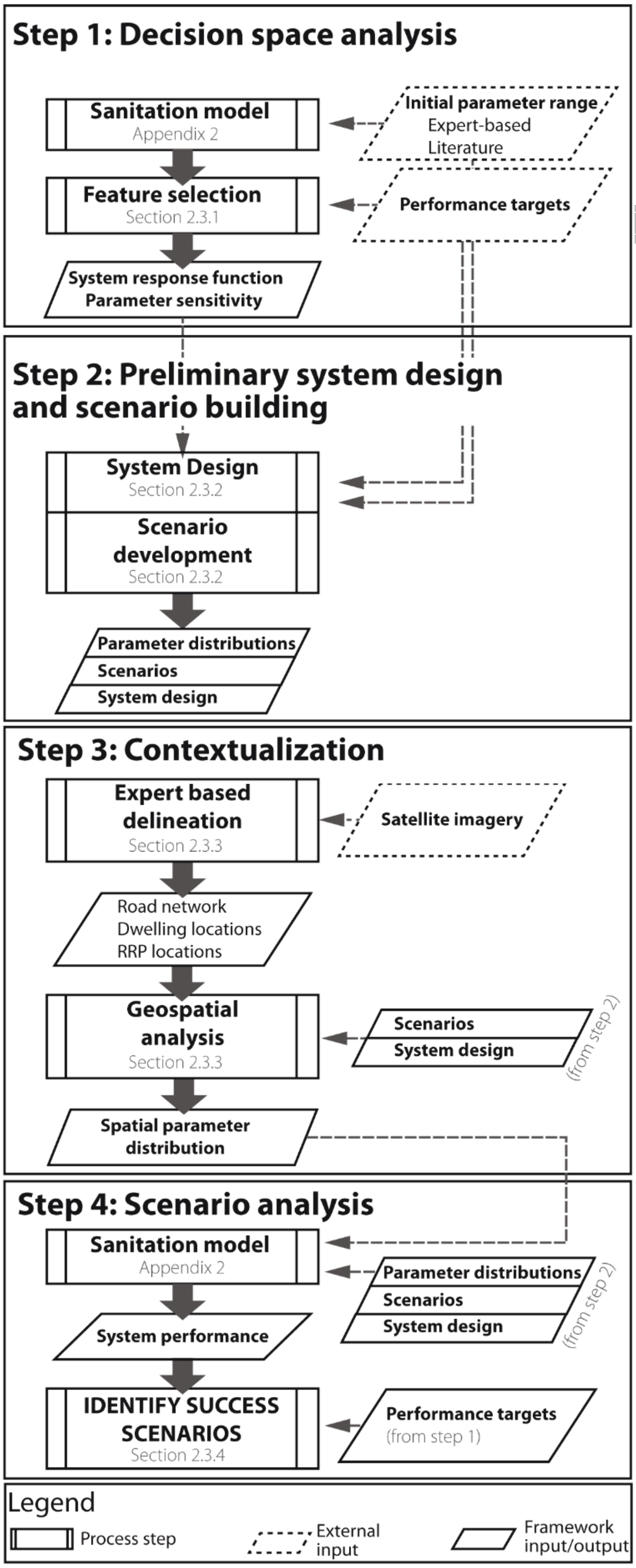

Figure 1. Steps of the proposed framework, with reference to relevant sections, and the respective inputs and outputs of each step.

2.1. Numerical model development

We developed a stochastic numerical model that dynamically simulates the filling of $1 \ldots n_{Fac}$ facilities and their emptying through a demand-driven sanitation service system (for a full list of variables see Appendix 1, for a detailed description of the model see Appendix 2). n_{Fac} is defined by the total system capacity (number of users connected), and the number of users per toilet facility. The model simulates a service vehicle traveling between facilities on service rounds. A service round begins and ends in the central treatment unit. During each service round, the service personnel travels between full facilities, empties them, and returns to the central treatment unit as soon as the capacity of the vehicle is reached. The number of rounds the vehicle can complete is limited by the length of a working day. The model results in two measures of system performance: 1) service capacity (number of users served) and 2) cost per user.

ACCEPTED MANUSCRIPT

2.2. Case Study

We demonstrate the application of the proposed framework for a sanitation system designed within the "Reinvent the Toilet Challenge" (RTTC) program initiated by the Bill and Melinda Gates Foundation. RTTC aimed to develop novel sanitation solutions for the urban poor based on some initial premises, mainly competitive costs below 0.05 \$/user/day and applicability in dense informal settlements. General details are available from http://www.bluediversiontoilet.com and (Larsen et al., 2015), technical aspect of toilet facility design are discussed in Künzle et al., (2015). User acceptance of the toilet facilities and service system was high under field conditions (Tobias et al., 2017).

Within the project, a household-level sanitation service based on source-separating toilet facilities connected to a semi-centralized treatment facility (resource recovery plant, RRP) by a corresponding transport system was designed based on the herein proposed framework. We defined a performance target for costs ($C_{tot,user}$, including transport and emptying) of below 0.015 \$/user/day (= 30% of total cost limit). The limited availability of space in dense informal urban settlements led to an initial technical design with shared, low-storage toilet facilities (each toilet facility consisting of two separate toilet interfaces; (Larsen et al., 2015)). We targeted 20 users (approx. 4 families) per toilet facility.

These facilities are equipped with exchangeable containers for feces that are exchanged on average twice a week for empty ones when full. Urine is stored in built-in containers emptied by electrical pumps. Both types of containers have a few days of storage capacity (on average 3.5 days), and both urine and feces are transported by a small service vehicle. The facilities are equipped with basic telemetry to signal their fill-level to the service provider at pre-defined times such that timely service can be scheduled. We define that the system fails if not all facilities that require service are serviced on the day that they are full. Such a failure should not occur on more than 5% of days (we selected 5 % as a typical threshold in engineering applications; Brown et al., 2012). Full facilities that are not serviced in time will then add to the service demand on the next day. Obviously, the call-for-service is activated before the toilet is full: if the signal is for instance given in the morning, the toilet must be usable during the entire day because service may only be provided in the evening. We do not explore the costs and benefits of any safety factors, neither do we model the use of continuous real-time information on the fill level. In reality such features could facilitate the planning process, e.g. by leading to premature emptying of 'nearly-full' facilities on days with otherwise little activity, but such refinements of the model are beyond the scope of this paper. The transport system should operate within the cost limit for small system capacities or low user densities (e.g. defined as user per hectare) to enable initial system implementation without subsidies and reduce financial risk. Based on these premises, we derived initial ranges for the system parameters and validated these parameters with sanitation practitioners during various workshops (for illustrative purposes, Table 1 already introduces the results of this consultation). Parameters are assumed to be uniformly distributed, because for many parameters, only minimum and maximum values were available in step 1.

ACCEPTED MANUSCRIPT

Table 1: Initial parameter distribution for a sanitation service system. Parameters are assumed to be uniformly distributed between minimum and maximum values. Modeling is performed for a service system with only one vehicle and one RRP per system. For the initial modeling in Step 1, one worker per vehicle is assumed.

Parameter		*Min*	*Max*	*Unit*
Cost of capital, interest	ι_{Cap}	3	6	*%*
Distance facility-facility	$d_{Fac\text{-}Fac}$	10	1000	*m*

Distance treatment - facility	$d_{Tr\text{-}Fac}$	10	1500	*m*
Facility holding capacity	$cap_{Fac,max}$	100	300	*kg*
Fraction of facilities on path	f_{Path}	0	100	*%*
Fuel consumption	r_{fuel}	0	4	*l/km*
Fuel price	p_{fuel}	1	2	*$/l*
distance on foot path	d_{Path}	0	0.15	*km*
Product accumulation rate (urine)	m_U	0.4	2	*l(urine)/user/d*
Product accumulation rate (feces)	m_F	0.2	0.5	*kg(feces)/user/d*
Payment, service worker	s	3	8	*$/worker/d*
Service time per facility	t_{Serv}	5	60	*min/facility*
Vehicle maximum transport capacity	$cap_{V,max}$	100	800	*kg*
Initial capital costs (vehicle and equipment)	$c_{C,0}$	250	3500	*$*
Vehicle speed on roads	v_{road}	1	5	*km/h*
Walking speed on paths	v_{path}	0.1	1	*km/h*
Working hours / day	t_{max}	4	12	*h/d*

ACCEPTED MANUSCRIPT

Three informal settlements with distinct spatial properties were selected as potential implementation sites. Settlement S1 (32°35'20,017"E, 0°20'56,66"N) and S2 (32°35'6,422"E, 0°21'7,435"N) are located in Kampala, Uganda, and settlement S3 (81°35'39,7"E, 21°13'30,048"N) in Raipur, India. They represent a sparsely populated, peri-urban settlement (S1), a dense, urban settlement (S2), and a situation where dense pockets of informal settlements are separated by areas of regular housing development (S3). A household survey is available for S1 and S2: it states an average household size close to five (Tumwebaze et al., 2014). The same household size was applied to S3.

2.3. Developing a robust planning framework for sanitation services

2.3.1. Step 1: Decision space analysis

For the analysis of the decision space, the model is run many times and the values of its input parameters are obtained from a Monte Carlo approach (Rubinstein and Kroese, 2011) according to the defined uniform distributions (Table 1). We identified those of the parameters shown in Table 1, to

ACCEPTED MANUSCRIPT

which system capacity and costs are most sensitive, i.e., which have the highest impact on system performance. Correlating the most sensitive parameters and the system performance allows low-dimensional, visually interpretable representations of system performance, known as system response functions (Brown et al., 2012), to be derived for both system costs and capacity. The system response functions also map critical thresholds of system performance within the decision space. For this paper, we ran 25,000 independent simulations and derived system response functions in two steps. First, we applied a sequential forward-feature selection algorithm (FSA) to identify the most sensitive system parameters. We selected this approach for its relative simplicity and computational efficiency (Saeys et al., 2007) as well as for its standard implementation in the modeling environment (The MathWorks, Inc., 2014). Within this algorithm, we implemented a neural network as a flexible, nonlinear explanatory model (May et al., 2008). Second, the system performance was mapped throughout the entire parameter space for the two most sensitive parameters. For this mapping, we selected an analytical formulation of the form

$$\mathbf{R_m} = \mathbf{a} * \boldsymbol{p_{1,m}} + \mathbf{b} * \boldsymbol{p_{2,m}} + \mathbf{c}, \qquad \textbf{EQ 1}$$

where p_1 and p_2 define the two first parameters selected by the FSA. The subscript m refers to any analyzed model output, i.e. cost or capacity. R_m can be evaluated analytically throughout the parameter space for p_1 and p_2. The response function can then be analyzed visually. From this visualization, we identify the ranges of p_1 and p_2 that result in a system performance matching the performance target.

Obviously, accuracy is lost during the process of converting the original model into the lower-dimensional response function (Brown et al., 2012). Additionally, a more detailed analysis of system sensitivity, by using the Sobol or similar methods (Saltelli et al., 2008), for instance, can be implemented in the future. However, this is not a major limitation at this point because the system response function only provides initial guidance towards preliminary system designs (i.e. the planner gets a rough idea of what is possible and what is not), while the actual system design is analyzed in more detail in step 4 using the full model.

ACCEPTED MANUSCRIPT

2.3.2. Step 2: Preliminary system design and scenario building

The development of a preliminary system design is a central step of the proposed methodology. Based on the system response functions found in Step 1, the planner defines the sensitive parameters within a range that is likely to lead to sufficient system performance. At the same time, more effort is spent on the collection of realistic information, especially on the sensitive parameters, in order to make sure that it is also possible to obtain a given performance in reality ('reality check'). This work can be based on common sense, literature, expert interviews, or even specifically designed field work and surveys. With this information in mind, the planner will define scenarios to be tested in more detail. These scenarios must help inform the business model and business plan of an entrepreneur wanting to start a sanitation-service business: which vehicle is preferable, how many workers must be hired, which payment scheme is suitable, how profitable is the business in the different stages of development, when will additional capital be required, etc.

The specific results of Step 2 are a more detailed estimate of the system design parameters, including uncertainty, and a set of relevant scenarios to be simulated.

2.3.3. Step 3: Contextualization

The aim of the contextualization step is to represent the spatial characteristics of a specific informal settlement in the stochastic model by measuring the empirical probability distributions of the spatial parameters defined in Appendix 1 ($d_{Fac-Fac}$, d_{Tr-Fac}, d_{Path}, and f_{path}) from remote sensing data. These parameters are relevant, a) because they directly impact system performance while in turn being b) a function of certain design decisions or variable between scenarios (e.g., users per facility and thus total facility number and distance between facilities)). Hence, step 3 takes up results from step 2, and uses them in a geo-spatial analysis based on free-of charge remote sensing data. Since the location of toilet facilities and treatment plants are unknown at this early planning stage, the geo-spatial analysis (referred to as contextualization) approach has to be probabilistic.

The analysis is performed in a standard Geo Information System (e.g. ESRI ArcGIS, or the (open-source) QGIS). Most GIS software allow to access high resolution satellite imagery, e.g., from Google

Maps, or Bing Maps (this analysis was based on imagery derived from Bing Maps in ArcGIS). With the increasing availability of satellite data, it is likely that such data are now available for nearly all geographic settings, and will remain free of charge in the foreseeable future. The satellite images are used to delineate the road network of the settlement and to identify dwellings and potential locations for the treatment plants (open areas > 200 m^2, close to major roads) in the settlement. The road network is classified visually either as footpaths (unusable by larger service vehicles) or larger roads. The result is a digital representation of the roads in a settlement as a network structure. Please note that informal settlements are characterized by very simple infrastructure and that even larger roads will mostly be unpaved and not subject to much traffic.

ACCEPTED MANUSCRIPT

For the interpretation of the satellite images, we assumed that each dwelling accommodates one household and is a potential location for a toilet facility (however, toilet facilities are to be shared amongst multiple households; see section 2). Candidate locations for toilet facility are selected as a subset of size l_{fac} of the dwellings, where l_{fac} is defined by

$$\boldsymbol{l_{fac}} = \frac{\boldsymbol{cap_{Sys}}}{\boldsymbol{n_{user,fac}}} \qquad \textbf{EQ 2}$$

e.g. if the system capacity is 200 users $(cap_{Sys} = 200)$ and one facility is shared among 20 users (see section 2, $n_{user,fac} = 20$), then ten of the mapped dwellings are randomly selected to represent a toilet facility. Next, we identify for each toilet facility the nearest point in the road network and determine if it is on a road, or on a footpath. On this basis, we calculate the fraction of toilet facilities on footpaths, f_{path}.

A shortest-path algorithm (Dijkstra, 1959, as implemented in ArcGIS 10.0) can then be used to measure the shortest route between any pair of system elements based on the digitized road network. For facilities that are assigned to a footpath, we use the same routing algorithm to determine the distance on that footpaths to the next road, d_{Path}. We also use the routing algorithm to measure the mean distance from each candidate site for the treatment plant to all accessible (i.e., located on roads) toilet facilities. We select that location for the treatment facility that minimizes the distance to all accessible facilities.

The application of this shortest path algorithm is based on the assumption that the worker is able to find a good route based on his/her experience.

ACCEPTED MANUSCRIPT

Measuring the distance distributions between toilet facilities is more complex, because it requires the service interval of the toilet facilities to be considered in the measuring procedure. This is because the facilities are not emptied daily, so it is not sufficient to measure the distance from each facility to its direct neighbor. When the service person has finished servicing a toilet facility, he/she has to travel to the next full toilet facility. However, there is a high probability that the next facility will not require servicing. This probability, P, is defined by

$$\boldsymbol{P = 1 - \frac{1}{i_{serv}}} \qquad \textbf{EQ 3}$$

where i_{serv} is the mean service frequency for the facilities. The probability that none of the next n_N toilet facilities requires service is approximated by:

$$\boldsymbol{P' = \left(1 - \frac{1}{i_{serv}}\right)^{n_N}} \qquad \textbf{EQ 4}$$

Where i_{serv} is the average service interval (in our system 3.5 days, see section 2), and n_N was determined such that P'_i< 0.05 (meaning that that there is a 95% probability to find a full facility within the nearest n_N facilities). Based on EQ 5, there is thus a 95 % probability that for our system, the next full facility will be found amongst the neighboring 9 facilities ($n_N = \mathbf{9}$). On that basis, the routing algorithm determines the travel distance ($i.e., d_{Fac-Fac}$) from each accessible toilet facility to its n_N neighbor facilities. This method allows us to consider the impact of service frequency on inter-facility distances: the higher the service frequency (which is defined in the system design step by setting the holding capacity and user number of facilities) and the more facilities are installed in a given area (which is defined by the scenarios, i.e., how many users are to be connected to the system), the lower will be the travel distances between facilities.

These procedures yield empirical distributions for $d_{fac-fac}$, d_{Tr-Fac}, and d_{path}. Empirical distributions are transformed to analytical ones to represent spatial characteristics within the stochastic system model. In this study, the selection of an analytical distribution that best represents the measured empirical distribution was performed according to the methodology proposed by (Sheppard, 2012). If the stochastic model is used to simulate systems of different sizes (i.e. the number of users connected to the system), the measurement of the spatial parameters must be repeated for each system capacity.

ACCEPTED MANUSCRIPT

The specific results of Step 3 are the optimal position of the treatment plant and stochastic formulations of the empirical distances between plant and facility and between full facilities, the number of facilities on paths, and the distances from facilities on paths to the next road.

2.3.4. Step 4: Scenario analysis

We parametrized the model according to the preliminary design defined in Step 2 and the spatial parameters derived in Step 3. The model simulates the system behavior for different scenarios for a large number of future conditions. For our case study, we ran the model for 15 years of continuous system operation (5,475 times) for each scenario. Values of model parameters that can change from day to day (e.g., excreta accumulation in each facility is very likely to vary from day to day) or from round to round, are re-sampled in each run or in each round. The model results identify scenarios for which the sanitation system matches the performance targets in terms of reliability and costs.

If the performance is insufficient for all scenarios, step 2 would be repeated: The system design must be changed and/or new scenarios developed. Step 4 should be repeated during system implementation when practical experience will allow the model parameters to be updated, thus enabling a continuous re-evaluation of long-term system performance under observed field conditions.

3. Results

Step 1: Decision space analysis

As first step, we performed a Monte Carlo simulation with 25000 runs. We applied the FSA to the results and identified those input parameters (i.e., all parameters listed in Table 1) that correlate most to system cost and capacity (i.e., to which cost and capacity are most sensitive).

Table **2** shows the results of this analysis, i.e., the cumulative $\boldsymbol{R^2}$ that is reached by adding additional input variables to the explanatory model in the FSA (see section 2.3.1). The FSA selected four variables for capacity and seven for costs and the final, cumulative $\boldsymbol{R^2}$ is above 0.9 for both capacity and costs. It should be noted that we considered system capacity in the analysis of system costs, even though system capacity is an output of the model and not an input variable. We did this because system capacity (i.e., how many users are connected to the system) strongly impacts cost per user through the split of fixed costs amongst all users.

ACCEPTED MANUSCRIPT

Table 2: Input parameters selected by the FSA and resulting cumulative R^2 for system capacity and costs

a) SYSTEM CAPACITY (cap_{Sys}) *[USERS]*

Added input parameter	**1. Maximum working hours/day**	**2. Service time per facility**	**3. Vehicle speed on roads**	**4. Vehicle capacity**
	t_{max}	t_{Serv}	v_{Roads}	$cap_{V,max}$
Cumulative R^2	0.579	0. 854	0.937	0.980

b) SYSTEM COSTS ($C_{tot,user}$) *[\$/USER/DAY]*

Added input parameter	**1. Fuel consump-tion**	**2. Distance facility-facility**	**3. System capacity**	**4. Fuel price**	**5. Distance RRP to facility**	**6. Vehicle capacity**	**7. Payment, service person**
	r_{fuel}	$d_{Fac-Fac}$	cap_{Sys}	p_{fuel}	d_{Tr-Fac}	$cap_{V,max}$	s_1
Cumulative R^2	0.376	0.685	0.753	0.820	0.862	0.896	0.934

Service capacity correlates strongly with operational (working hours) and technical (service time per facility) parameters. Vehicle parameters (i.e. speed or capacity) are of less importance. Costs correlate above all with fuel consumption and inter-facility distances. System capacity, i.e. how many users are connected within a given area, was selected as a third parameter.

Based on Table 2, graphical system response functions were generated from EQ 1 to illustrate the dependence of system capacity and costs on the first two variables selected in each case (Figure 3a,b). For system costs, a second response function was derived (Figure 3c) considering the second and third selected variables (inter-facility distances, system capacity), in order to qualitatively evaluate the impact of different strategies for system up-scaling on the costs per user.

ACCEPTED MANUSCRIPT

In accordance with the results of the feature ranking (Table 2), the response function for system capacity (Figure 3a) has a higher correlation (R^2=0.85) with the model results than the cost response function (Figure 3b; R^2=0.65). The second cost response function reaches only a low R^2 (0.35). This response function represents general trends rather than giving an accurate picture of the system behavior (Figure 3c).

TED MANUSCRIPT

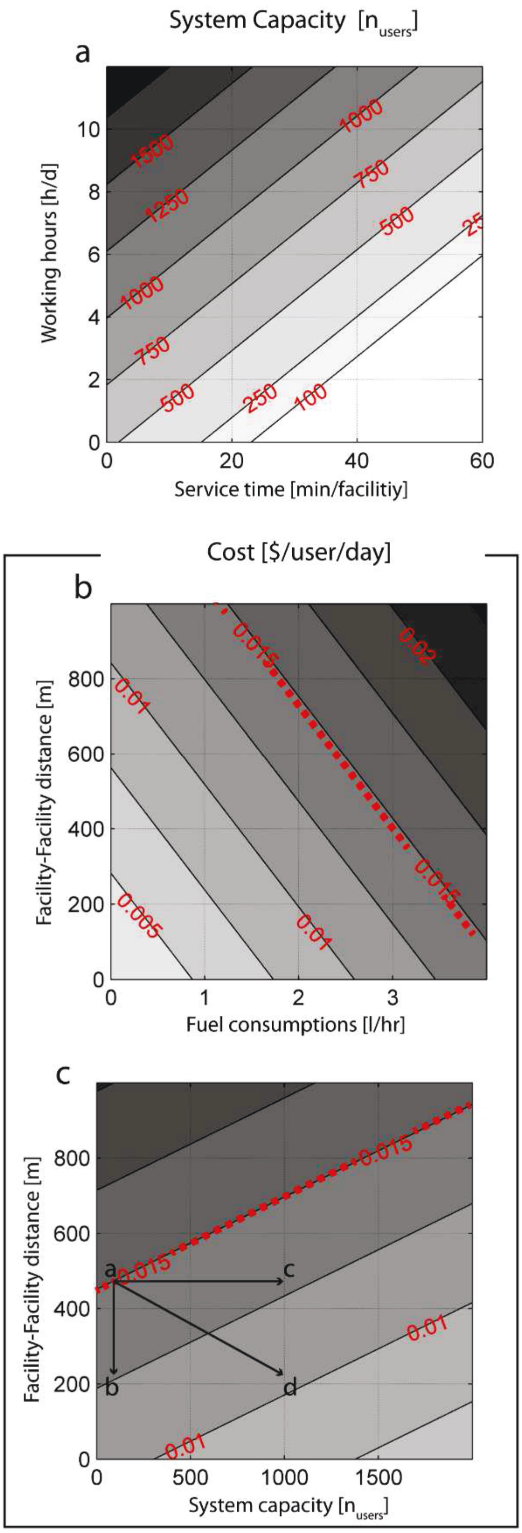

Figure 2: System response functions for capacity (a) and cost (b and c). To assess scale effects, a second cost response function (c) was derived. The dashed red line in b and b indicates the cost limit. Lower-case letters and arrows in c refer to different expansion strategies and their impact on costs as discussed in the text.

System capacities up to 2000 users (Figure 3a) are attainable with one vehicle and one service person per vehicle and the costs are below the threshold value (< 0.015$/p/d) for a wide range of system conditions (Figure 3b). Nevertheless, high system capacities require long working hours (> 8 hours/day) and very rapid (<5 min/facility) servicing of facilities. With eight working hours per day available and reasonably fast (15-20 min) emptying of facilities, the system can reach a capacity of 1250 users. A system with such fast-emptying facilities, but a lower user number does not require a full-time worker. For instance, a system for 750 users and with facilities that can be emptied in 20 minutes requires only 5 hours of daily labor. In turn, a system for which the service time of facilities is longer (~40 min) can still reach a capacity of 750 users if one full-time (i.e., 8 hours/day) worker is available. The cost target (0.015 $/user/day) is not exceeded for a wide range of system conditions, indicating high flexibility with respect to vehicle selection, spatial characteristics, and user number of the servicing system (Figure 3b). For systems with short distances between facilities (< 100 m) and a relatively fuel-efficient vehicle (< 1 L/h), servicing can be provided at around one third of the maximum cost. The second cost response function (Figure 2c) demonstrates how changing system capacity and the spatial distribution of users (i.e., distances between facilities) impact costs. In comparison to a system that just meets the cost limit (Figure 2c, point a), costs could be decreased by decreasing the distances between facilities, i.e., by connecting the same user number within a smaller area (moving from point a to b in Figure 2c). Cost can also be decreased by increasing the system capacity. System capacity can be increased through two strategies, either by spatial expansion, i.e., increasing system capacity without decreasing travel distances (moving from point a to c in Figure 2c), or by densification, i.e., increasing system capacity by including new users within a given area (moving from point a to d in Figure 2c). It is evident from Figure 2c that densification has the highest potential to reduce costs per user.

ACCEPTED MANUSCRIPT

Step 2: Preliminary system design and scenario building

On the basis of the findings in step 1, we designed a system with input parameters that would have a high chance of leading to a successful transport system. We targeted a system based on a single service vehicle and a service time of around 25 minutes, aiming at a system capacity of 700 to 1000 users (Figure 3a). We focused the scenario development on vehicle selection, payment schemes and upscaling of the system (i.e., increasing system capacity), all parameters of high importance for the financial viability of the transport system. We considered only vehicles commonly used in informal settlements, specifically a manual pushcart (requiring two workers) and a motorized two-wheel tractor (requiring only one worker) with a detachable trailer (Coffey and Coad, 2011). We again assumed that there is only one vehicle and one RRP per system. Each vehicle needs to be equipped with a small wheelbarrow to access narrow footpaths. The capacity of such a wheelbarrow is assumed to be around 200 kg, defining the upper limit for the holding capacity of the facilities (Larsen et al., 2015). We represented the uncertainty in all parameters by using a normal distribution with $\sigma = 0.1 \cdot \mu$, as no empirical data on the distributions of these parameters were available.

ACCEPTED MANUSCRIPT

We included ten different system capacities in the scenario analysis. The area covered by the sanitation service is fixed, i.e. users are distributed in the entire area for low system capacities, and the area is not expanded for high system capacities (i.e., densification is simulated). In total, we analyzed 120 scenarios (2 vehicles*2 payment schemes*3settings*10 user densities). For each scenario, we evaluated the cost per user, the actual capacity (i.e. how many facilities are serviced in a day?) as well as the service demand (i.e. how many facilities need to be serviced in a day for a given system capacity?). We calculated the failure probability by comparing service capacity and service demand on each day, i.e. this probability measures the number of days during which service demand exceeds service capacity. The model considers that unserved facilities add to the next day's service demand. Hence, unserved facilities accumulate in the system if service demand constantly exceeds service capacity. A system that can service unserved facilities from the previous day on the next day can display an increased failure probability, but the average service demand and capacity should not diverge.

Table 3: Parameter distribution for the proposed sanitation service system resulting from the preliminary system design process. Distributions of spatial parameters will be derived empirically for selected case study sides in Step 3 and are thus left undefined here.

Parameter		μ	σ	*Unit*	*Reference*
Cost, Labor					
Payment scheme 1	s_1	0.5		*$/facility*	#
Payment scheme 2	s_2	5		*$/worker/day*	
Distance betw. facilities	$d_{Fac\text{-}Fac}$	Empirical		m	*
Distance RRP facility	$d_{Tr\text{-}Fac}$	Empirical		*m*	*

Facility holding capacity	$cap_{Fac,max}$	200		*kg*	1
Fraction of facilities on path	f_{path}	Empirical		*%*	*
Fuel demand	r_{fuel}				2
2-wheel tractor		2	0.2	*l/h*	
Pushcart					
Fuel price	p_{Fuel}	1.6	0.16	*$/l*	3
Length of path	d_{Path}	Empirical		*km*	*
Product accumulation	m_U	1.1	0.11	*l(urine)/user/d*	4
Product accumulation	m_F	0.23	0.023	*kg(feces)/user/d*	4
Service time	t_{Serv}	25	2.5	*min/facility*	1
User number	n_{user}	100-1000		*user*	1
Vehicle capacity	$cap_{v,max}$				2
2-wheel tractor		600	60	*kg*	
Pushcart		300	30	*kg*	
Vehicle maximum speed	v_{Road}				#
2-wheel tractor		5	0.5	*km/h*	
Pushcart		2	0.2	*km/h*	
Speed on paths	v_{Path}	0.3	0.03	*km/h*	#
Working hours	t_{max}	8	0.8	*h/d*	#

[1] Larsen et al. (2015); [2] (Coffey and Coad, 2011); [3] World Bank (2015); [4] Porto and Steinfeld 2000; Schouw et al. 2002; [#] Field interviews (Kampala, Uganda), or expert based; * from contextualization (Section 4, Step 3)

ACCEPTED MANUSCRIPT

Step 3: Contextualization

Figure 3 shows satellite images of the three settings and clarifies their various spatial characteristics. Table 4 summarizes the basic spatial characteristics of each setting. Key findings are that increasing facility density and user number effectively shortens travel distances between facilities independently of settlement structure and road infrastructure. This is of relevance for the sanitation-service system because the results in Step 1 indicate that both user number and inter-facility distances affect the costs (see Figure 3c). Increasing user density thus acts on two central controls and has a major potential to decrease costs per user.

For S1, there are multiple candidate RRP positions and a low fraction of dwellings on footpaths. Distances are longer than for the other settings because the total settlement area is larger. S2, in contrast, is characterized by a small settlement area, a high settlement density, and poor road connections. There is a single candidate RRP position, and many dwellings are located on footpaths. However, distances between dwellings on paths and the next road are relatively short. The spatial characteristics of S3 represent conditions between S1 and 2 with regard to all parameters.

We measured the distances between facilities ($d_{Fac-Fac}$) and between treatment plants and facilities (d_{Tr-Fac}) for system capacities of 100 to 1000 in steps of 100 users (equivalent to 5 to 50 facilities in steps of 5 facilities). We report the results in terms of user density for a better comparison between settings (Figure 4). RRP-facility distances (Figure 4a) are longest for S3 and shortest for S1. RRP-facility distances do not decrease with increasing facility numbers as new facilities are added at random locations throughout the settlement. Increasing the system capacity from 100 to 1000 users results in an average reduction of 74% in inter-facility distances (Figure 4b). In addition, the variability of these distances decreases strongly with increasing user numbers.

ACCEPTED MANUSCRIPT

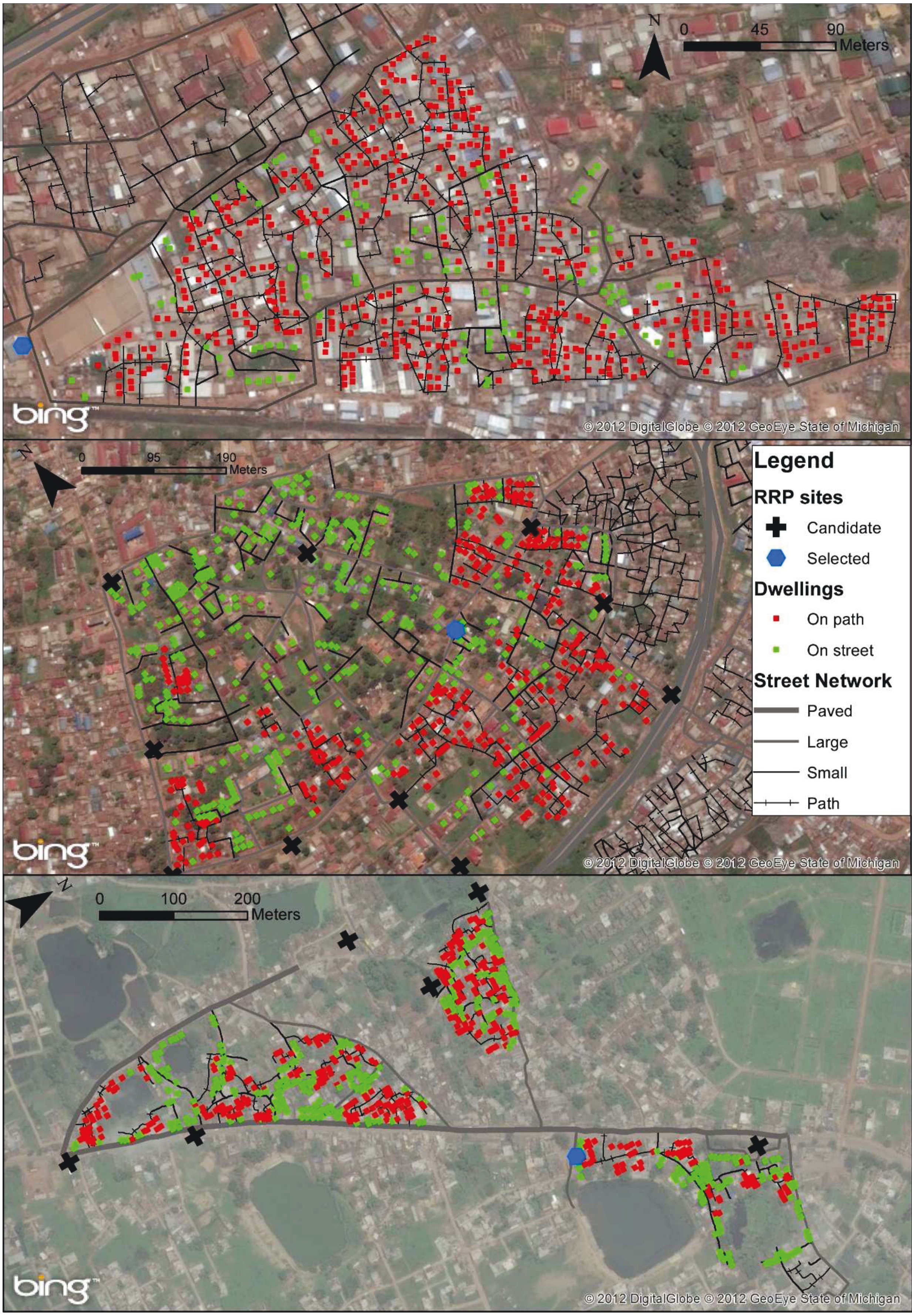
N
0 45 90 Meters
0 95 190 Meters
0 100 200 Meters
Legend
RRP sites
Candidate
Selected
Dwellings
On path
On street
Street Network
Paved
Large
Small
Path
bing
© 2012 DigitalGlobe © 2012 GeoEye State of Michigan
bing
© 2012 DigitalGlobe © 2012 GeoEye State of Michigan
bing
© 2012 DigitalGlobe © 2012 GeoEye State of Michigan

Figure 3: Satellite image of the three settings, including households, the potential and optimal RRP position and road infrastructure.

Table 4: Characteristics of three selected study areas in terms of population density, spatial setup and accessibility to households.

ACCEPTED MANUSCRIPT

	Settlement area [ha]	*Total Population [p]*	*Population density [p ha^{-1}]*	*Candidate RRP positions*	*% of houses on paths*	*Path length Mean/Max [m]*
S1	28.9	3700	128	9	46	34.1/177
S2	6.0	3120	520	1	83	26.6/68
S3	9.8	3163	320	7	55	21.4/81

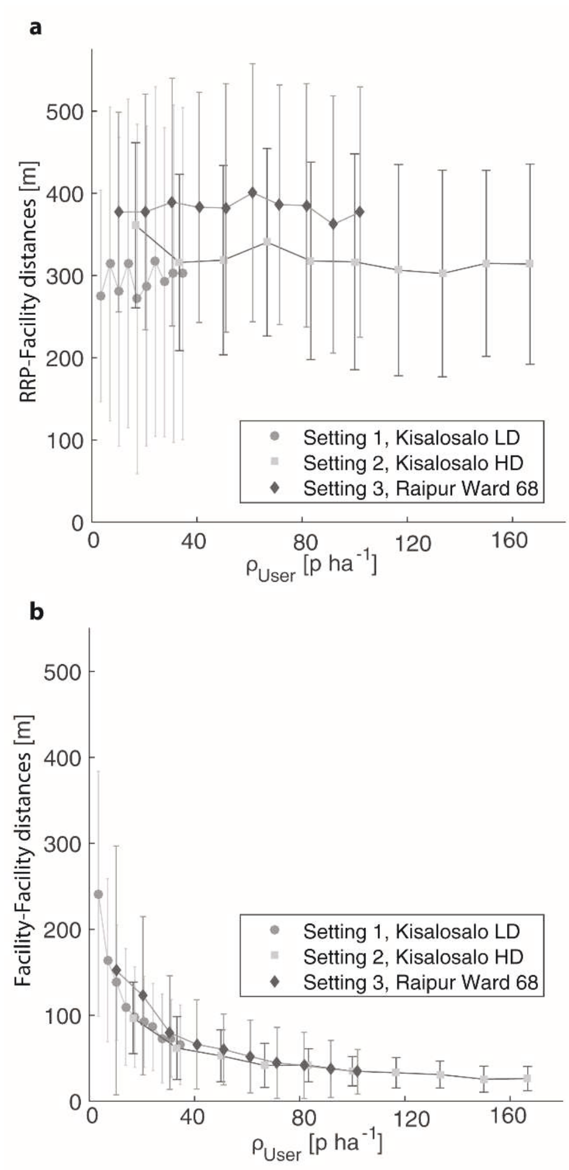

Figure 4: Correlation between increasing user densities (i.e. user per hectare) and required travel distances between RRP and facilities (a) and in between facilities (b) for the three settings shown Figure 3.

Step 4: Scenario analysis

We evaluated the performance for 15 years (5475 runs) of each of the 120 scenarios in terms of service capacity (limit: service capacity < service demand at less than 5% of runs) and costs per user (limit: 0.015 $/user/day).

ACCEPTED MANUSCRIPT

System Capacity

The pushcart-based service system is suitable for systems with a capacity of up to 400 – 500 users. It is evident from Figure 5 (left y-axis) how the number of serviced facilities, N (boxplots in Figure 5), increases with increasing system capacities. However, N reaches a limit at 12 (S3), 11 (S1) and 10 (S2) facilities, translating into a system capacity of 840 (S3), 770 (S1) and 700 (S2) users (20 users per facility). This means that a pushcart-based service system is suitable for up to 700 – 840 users. For higher user numbers, the number of serviced facilities, N, and service demand, D (triangular markers in Figure 5), start to diverge, indicating that full facilities accumulate in the system (see cutout in Figure 5). However, the failure probability already exceeds 5% for much lower system capacities (Figure 5, circle markers, right y-axis): 400 users in S1 and S2, and 500 users in S3.

The motorized service system is suitable for up to 700 – 800 users (Figure 6). N increases with system capacity, reaching a maximum of 14 facilities (corresponding to 840 users) per day independent of the setting (

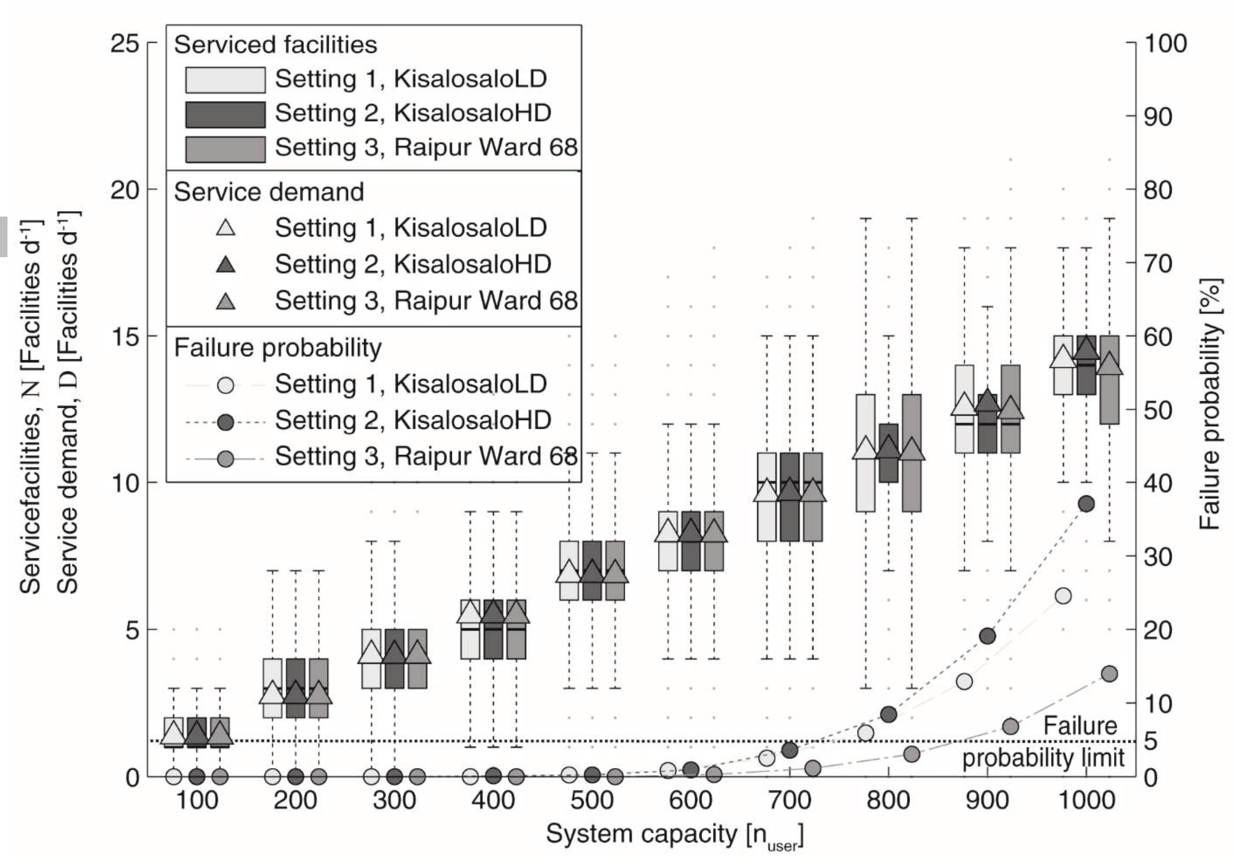

Figure 6). The failure probability is close to 0 for system capacities between 100 and 600 users and exceeds the 5% threshold only above system capacities of 700 (S1 and S2) and 800 users (S3), respectively. By comparing service demand D (triangle markers in Figure 6) and the median of serviced facilities, N (box plot in Figure 6) it is evident that there is only a small (< 0.5 facilities/day) divergence between service demand and the number of serviced facilities even for large system capacities. Hence, the motorized service system cannot guarantee on-time service in 15% (S3) to 35% (S2) of the time for the maximum system capacity, but it could service unserved facilities on subsequent days.

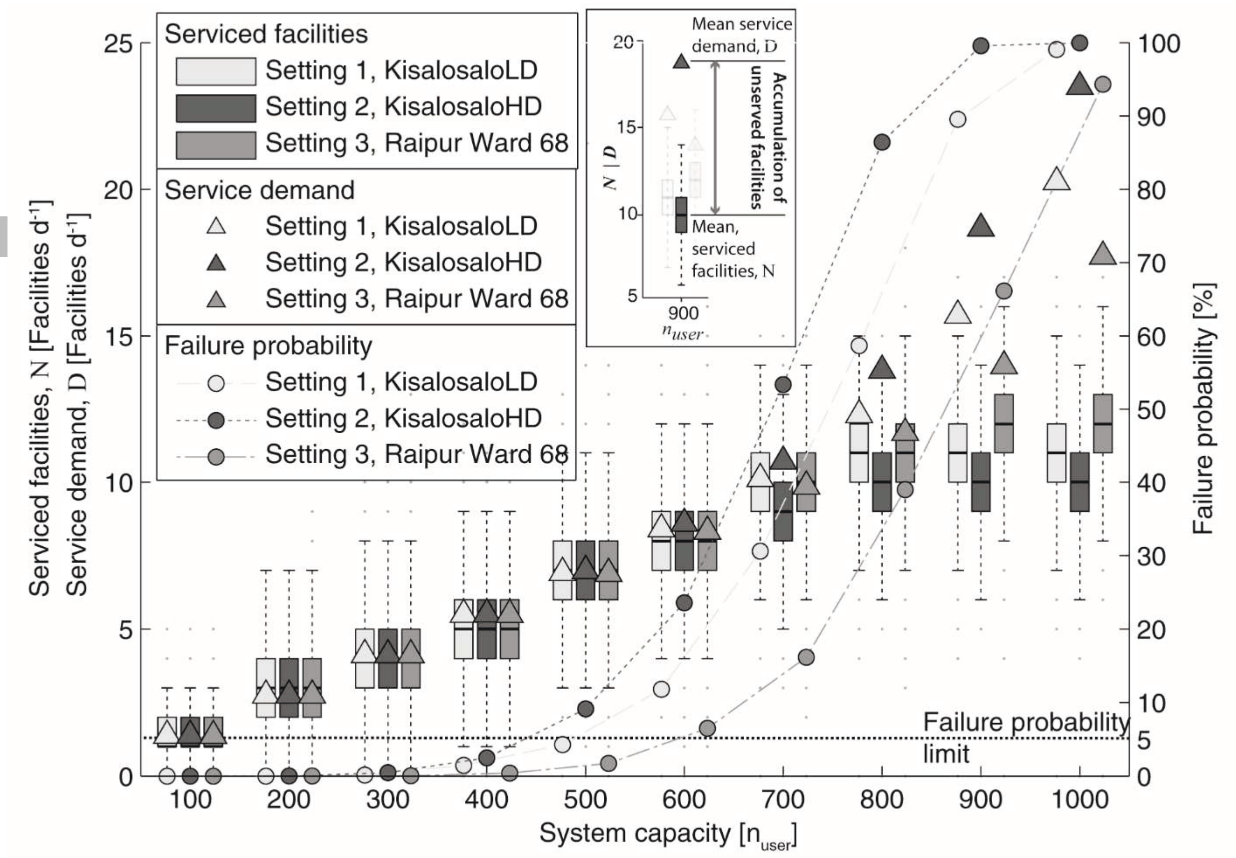

Figure 5: Serviced facilities and service demand (left y-axis) , and the resulting failure probability (right y-axis) for a pushcart-based service system for 100 – 1000 users. The cutout clarifies the divergence between service demand and service capacity, indicating the accumulation of unserved facilities in the system.

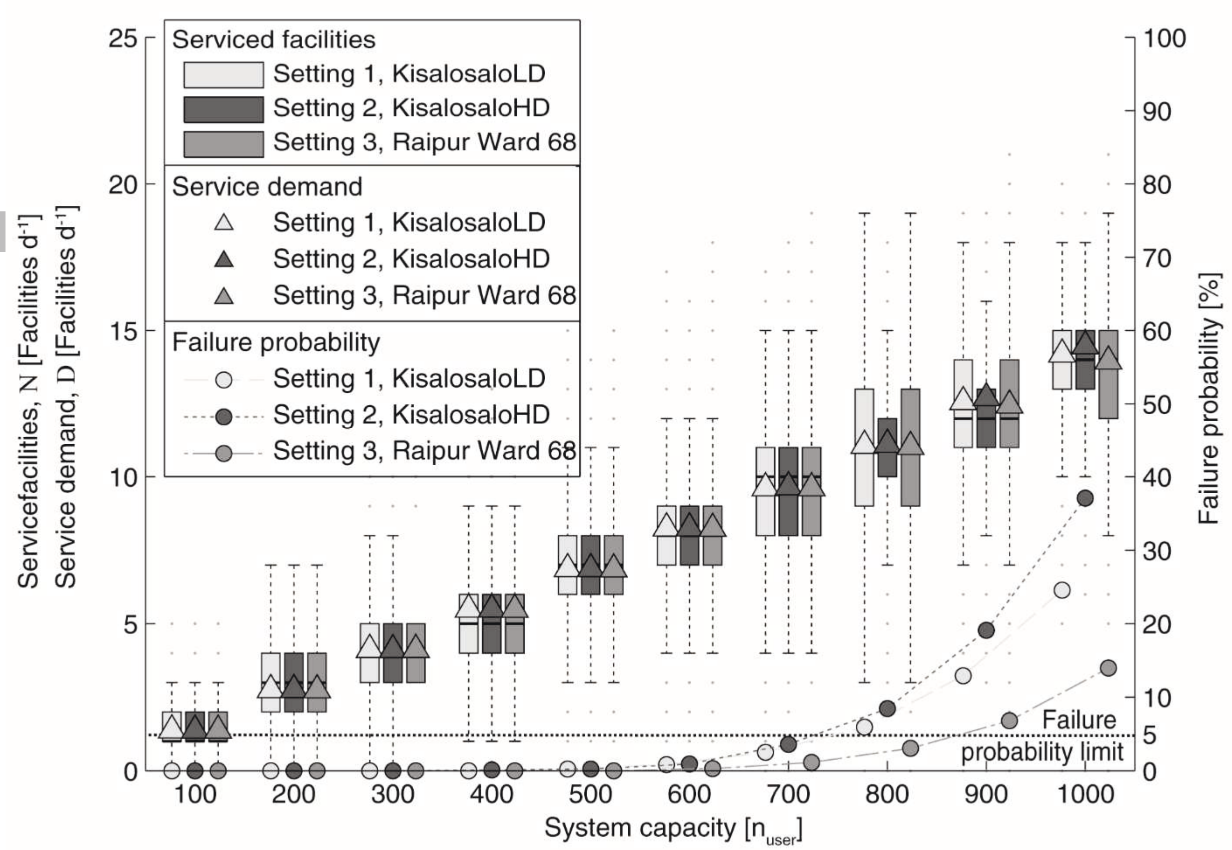

Figure 6: Serviced facilities and service demand (left y-axis) , and the resulting failure probability (right y-axis) for a motorized service system for 100 – 1000 users.

System Cost

The median costs of a pushcart-based service system with performance-based payment scheme (PS 1) meet the cost limit for all system capacities and settings (Figure 7a). The 95% confidence interval of costs is below the cost limit for all settings above a system capacity of 200 users. The pushcart-based service system with fixed daily payment scheme (PS 2) is not economically viable below system capacities of 700 users (Figure 7a). This is due to high labor costs (two service workers per vehicle). The costs for the two payment schemes converge at higher system capacities, when the fixed costs of PS 2 are split amongst an increasing number of users.

The median costs of a motorized service system match the cost limit for PS 1 from system capacities of 300 (S2) and 400 users (S1 and S3) respectively (Figure 7b). From system capacities of 500 (S2) and 600 users (S1 and S3) respectively, the 95% confidence interval of costs also falls below the cost limit. The motorized service system with PS 2 requires a minimum system capacity of 600 to 700 users (Figure 7b) to match the cost frame. For motorized systems, the median costs for all system capacities are different between settings. This is because of the fuel demand that links travel distances to costs, and hence to different spatial characteristics of settlements. The lowest costs occur in S2, where distances are shortest (see Table 4). Nevertheless, the resulting difference between the settings is not significant. Variability in costs decreases with increasing system capacity, which is in line with the decreasing variability in inter-facility travel distances for higher user densities (Figure 5).

ACCEPTED MANUSCRIPT

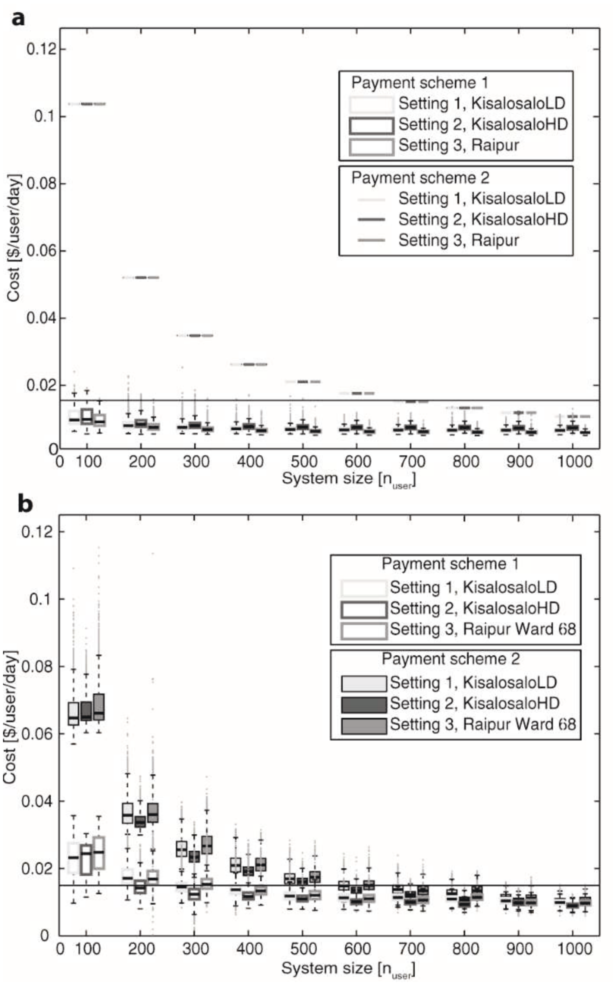

Figure 7: Cost incurred for the sanitation service using a pushcart (a) or a two-wheel tractor (b) and PS 1 and 2. The solid line indicates the cost limit.

Conditions for successful system implementation

The previous analyses indicate a strong trade-off between costs and failure probability. Increasing system capacity reduces system costs (Figure 8) but increases failure probability, especially for the manual service system (Figure 5). By analyzing the tradeoffs between cost, system capacity and failure probability, we identify conditions under which the system attains the performance targets. We report aggregated results over the three settings in terms of median cost and failure probability, and differentiate between two performance levels (PL): (1) performance above the target (i.e., < 5 % failure probability and < 0.015$/user/d) on 95% of the days, (2) performance above target on 50% of the days.

ACCEPTED MANUSCRIPT

Under PS 1, a pushcart-based transport system meets PL1 up to 400 users, above which the failure probability rises so rapidly that there is no interval where only PL2 is met (Figure a). The two-wheel tractor meets PL2 from 300 users, and PL1 for 500-700 users (Figure b).

Under PS 2, a pushcart-based transport system does not meet PL1 or PL2 (Figure c). A motorized system meets PL2 only for a single system capacity (600 users), and never PL1.

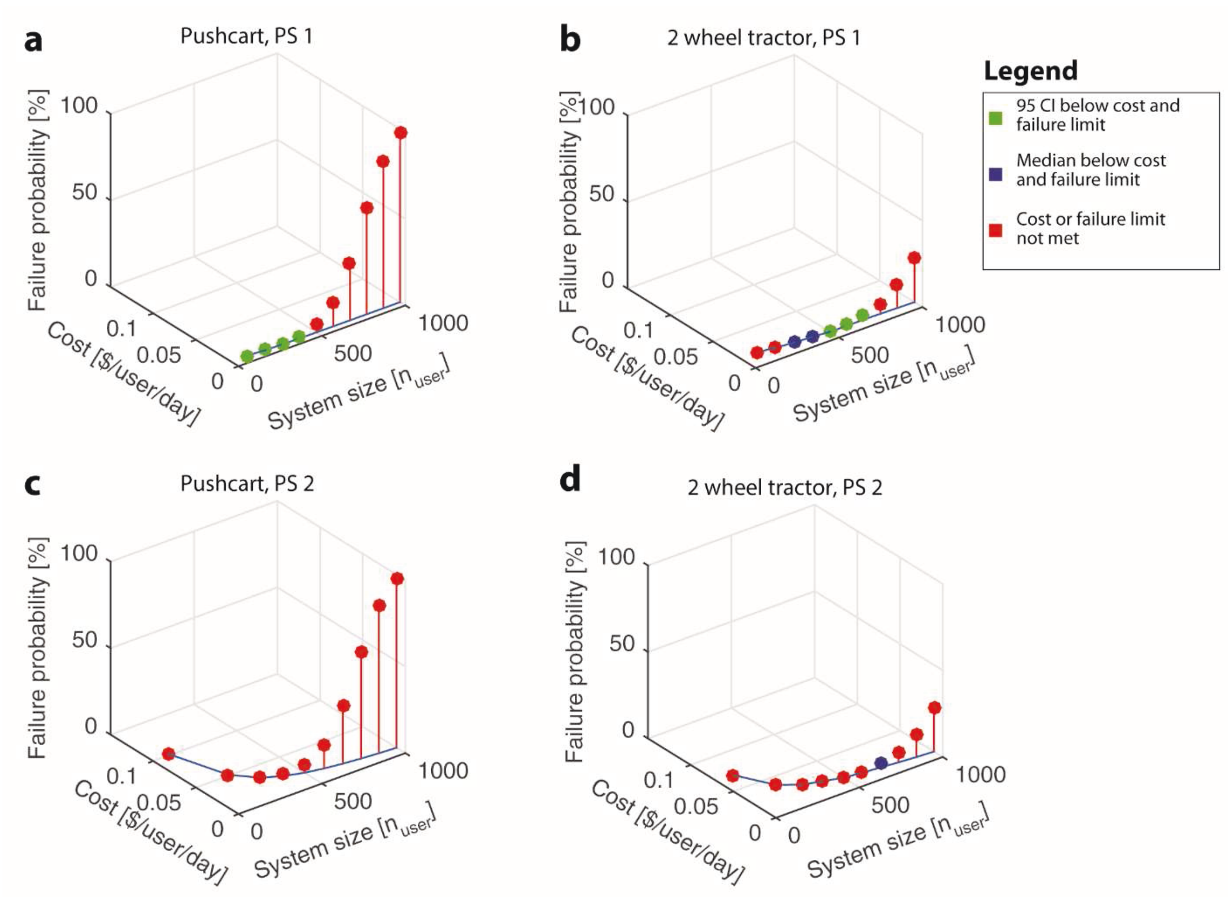

Figure 8: Tradeoffs between costs and system capacity for different user numbers, payment schemes (PS) (PS 1 in a and b, PS 2 in c and d), and transport solutions (pushcart in a and c, 2 wheel tractor in b and d).

4. Discussion

ACCEPTED MANUSCRIPT

Demand-driven service systems have the potential to improve sanitary conditions in informal settlements. With this paper, we present a first example of how such a demand-driven sanitation service system can be designed on the basis of a quantitatively robust planning framework, rather than with *ad-hoc* planning.

4.1.Critical parameters for successful sanitation services

The presented modeling is based on rough parameters and simplified assumptions. However, the results still identify general critical features in transport-based sanitation service systems. Such systems require toilet facilities that are fast to service and highly efficient workers, they apply strict cost management (as exemplified by the different payment schemes), and if possible, they have a flexible work force.

Obviously, fast-to-service toilet facilities contribute to high work efficiency. When designing the toilets, important factors are thus the accessibility of the feces container and a simple feature for pumping urine. The objective is to make emptying fast and less tedious for the workers. Especially for sustaining hard physical work over a longer time period, the ergonomics of the working situation is extremely important (Halim et al., 2014). Since informal settlements are often situated in countries with high temperatures in summer, the heat effect on work efficiency has also to be taken into account (Lundgren et al., 2014). Apart from being essential for the economics of transport logistics, the payment scheme based on the number of serviced facilities may also have an effect on work productivity. Despite methodological challenges, some authors were able to empirically demonstrate the positive effect of performance-based payment on work productivity (Oah and Lee, 2011). The advantage of a flexible workforce arises from the principle of service-on-demand and the intrinsic variability of toilet filling, but will have physical limits based on the length of the day and possibly also on the fact that heat may be prohibitive for hard physical work in the middle of the day.

Vehicle selection has a major impact on the performance of the service system. The results shown in Figure indicate that a pushcart is the only feasible solution for smaller system capacities (i.e. below 400 users) because of the high costs of motorized systems. For higher user numbers, a second pushcart, which would double the system capacity, could be deployed to avoid common challenges relating to motorized vehicles (e.g. maintenance cost and skills and the availability of spare parts (Coffey and Coad, 2011)). A motorized vehicle can be considered in specific cases, particularly if it proves difficult to find a qualified workforce to man additional vehicles. It has to be taken into account that whereas a pushcart demands harder physical work, two workers will be available for emptying the facilities – an effect that we did not consider in this article.

ACCEPTED MANUSCRIPT

4.2. Robust planning of sanitation services

We show that under certain conditions, regular and reliable logistics of human waste is feasible in informal settlements even within an ambitious cost frame and over a range of system scales. In general, this paper also provides insights into the drivers behind the performance and sensitivity of demand-driven service systems in informal settlements, making the results of general interest for service planning in informal settlements. With regard to system sensitivity, the model results agree with field observations (Coffey and Coad, 2011) in the sense that they identify the importance of work productivity and vehicle selection. Increasing system capacity (i.e. number of connected users), especially combined with densification of the facilities, has several positive effects on costs, mainly because of economies of scale and because it reduces distances. Nevertheless, our results point out that system expansion involves a major risk because of the strong trade-off between reliability and system capacity. This requires either a shift to a motorized transport solution or the purchase of an additional pushcart at a given point of expansion. This will, in turn, require sufficient reserves for financing the additional vehicle and training service personal. As Lempert et al. (2006) generically point out, “robust strategies are often adaptive” and similar findings hold for sanitation services. Our results indicate that adding some extra storage, which would allow for a delayed service, and more flexible working hours (i.e. work continues until all facilities are serviced on days with high service demand) could further increase system reliability and capacity. However, more detailed modeling is necessary to assess the optimal combination of flexible working hours and storage size.

Although we applied the framework to the collection of human excreta, it can be similarly applied to waste collection or the distribution of consumer and health goods in informal settlements. The development of the framework was strongly motivated by the rise of novel planning strategies that encourage quantitative planning approaches and numerical models to make use of uncertain but potentially useful information, such as from remote sensing. Similar to Brown et al. (2012), we base this framework on an inverted design process that begins with the identification of general thresholds, and their model-based mapping into the parameter space. These results then guide the further design steps, with each step being validated by the numerical model.

ACCEPTED MANUSCRIPT

4.3. Future developments and potentials

This approach can be improved in future, for example by applying spatial parameter estimation or using spatially explicit scenarios (Urich and Rauch, 2014). The expert-based delineation procedure is potentially subject to measurement biases, requires a certain expertise and is based on some assumptions (e.g. number of households per dwelling, number of persons per household). Such limitations could be overcome by automated mapping from remote sensing data and crowd-mapping of informal settlements (Kohli et al., 2012; Mattioli, 2014). Our proposed approach helped us to prioritize parameters that should be studied in more detail in the field (e.g. working habits). Also, additional system dynamics could be included in the model, e.g. the effect of multiple workers on the service time (see above), or using on-line information to optimize the service route (e.g. emptying nearly-full facilities on days with otherwise little activity). It should be kept in mind that the modeling approach presented in this study is not a stand-alone tool, but has many links to participative sanitation approaches. Implementing such a robust approach in wider planning frameworks can support more accurate model parameterization while helping to structure the various stages of a broader sanitation planning process, or to sensitize stakeholders and planners to system criticalities (Luethi et al., 2011; Tilley et al., 2014). During the implementation, parameter uncertainty can be continuously reduced by updating the model with new, better parameters estimates. The model can then be re-run and it can be re-evaluated if the reduction of uncertainty has an impact on the functioning of the service system and would imply any adaptive measures.

We thus hope that this paper will provide evidence for how numerical modeling and robust approaches can improve the ability to plan and analyze services for informal settlements in situations where typical top-down engineering approaches are not applicable due to the large uncertainty, and a reliance on *ad-hoc* self-organization often results in poor quality of services.

5. Conclusion

ACCEPTED MANUSCRIPT

This paper presents the development of a framework for the planning and analyses of service-based sanitation systems in informal settlements through four distinct steps. The framework is based on the application of a stochastic model of the sanitation system and a probabilistic analysis of remote sensing data. The level of detail evolves during the design process, from an initial generic assessment (step 1) to the analysis of a specific sanitation service system (step 2), its contextualization for specific informal settlement (3) and an analysis of system performance for numerous scenarios and system scales (step 4). The planning framework is robust in the sense that it bases the design of sanitation systems within a generic, quantitative risk analysis. The resulting design of the system is subsequently tested for a wide range of scenarios and system scales. The planning framework is applied to the design of a novel, transport-based sanitation service with toilets shared on the household-level. We required the system to work economically (below 1.5 c/user/day) with a single vehicle and for as little as 100 users (5 facilities). We find that such a system is feasible and highly scalable, in the sense that it can provide service from 100 (5 facilities) up to 700 users (35 facilities). We identify a strong trade-off between costs per user, as larger systems can provide the sanitation service cheaper, and reliability of the system, as more users also imply a higher probability of failure. Independent of the system capacity, work productivity and facilities that are fast to empty are most relevant. These two factors ensure that a high number of users can be serviced reliably, which then results in lower costs per user. Performance based payment, i.e., workers that are paid per serviced facility, and density of facilities, are other key factors to ensure financial viability and scalability of the system. The example given in this paper demonstrates that road-based services can provide a high level of service quality for sanitation in informal settlements under a wide range of operating conditions, spatial settings, and system scales. However, there is also a substantial risk for failure as sanitation services are sensitive to internal and external factors. We suggest that adaptive planning supported by quantitative frameworks, such as presented in this paper, will support service planning in informal settlements and reduce the risk of failure, even with limited field data and in conditions where nearly all parameters are deeply uncertain at the beginning of the planning process.

6. Acknowledgements

ACCEPTED MANUSCRIPT

The authors gratefully acknowledge Dr. Christoph Lüthi for suggesting the initial idea of a contextualized logistics model, and Prof. Dr. Hans-Joachim Mosler for his early support in focusing the scope of the paper. Ulrike Messmer's continuous inputs with regard to sanitation logistic chains greatly helped us to develop the scenarios and initial system designs. The authors also thank Andreas Scheidegger for his review of modeling and statistical aspects.

7. References

Brown, C., Ghile, Y., Laverty, M., Li, K., 2012. Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour. Res. 48, W09537. doi:10.1029/2011WR011212

Coffey, M., Coad, A., 2011. Collection of municipal solid waste in developing countries. UN Habitat, Nairobi.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271. doi:10.1007/BF01386390

Gates, M., 2010. What nonprofits can learn from Coca-Cola [WWW Document]. URL https://www.ted.com/talks/melinda_french_gates_what_nonprofits_can_learn_from_coca_cola/transcript (accessed 4.16.15).

Halim, I., Rawaida, Kamat, S.R., Rohana, A., Saptari, A., Shahrizan, M., 2014. Analysis of Muscle Activity Using Surface Electromyography for Muscle Performance in Manual Lifting Task. Applied Mechanics and Materials 564, 644–649. doi:10.4028/www.scientific.net/AMM.564.644

Hallegatte, S., 2009. Strategies to adapt to an uncertain climate change. Global Environmental Change, Traditional Peoples and Climate Change 19, 240–247. doi:10.1016/j.gloenvcha.2008.12.003

Kariuki, M., Jordan Schwartz, 2005. Small-Scale Private Service Providers Of Water Supply And Electricity : A Review Of Incidence, Structure, Pricing, And Operating Characteristics, World Bank Policy Research Working Papers.

Katukiza, A.Y., Ronteltap, M., Niwagaba, C.B., Foppen, J.W.A., Kansiime, F., Lens, P.N.L., 2012. Sustainable sanitation technology options for urban slums. Biotechnology Advances 30, 964–978. doi:10.1016/j.biotechadv.2012.02.007

Kohli, D., Sliuzas, R., Kerle, N., Stein, A., 2012. An ontology of slums for image-based classification. Computers, Environment and Urban Systems 36, 154–163.

Künzle, R., Pronk, W., Morgenroth, E., Larsen, T.A., 2015. An energy-efficient membrane bioreactor for on-site treatment and recovery of wastewater. Journal of Water Sanitation and Hygiene for Development 5, 448–455. doi:10.2166/washdev.2015.116

Langenhoven, B., Dyssel, M., 2007. The recycling industry and subsistence waste collectors: A case study of Mitchell's plain. Urban Forum 18, 114–132. doi:10.1007/BF02681233

Larsen, T.A., Gebauer, H., Gründl, H., Künzle, R., Lüthi, C., Messmer, U., Morgenroth, E., Niwagaba, C.B., Ranner, B., 2015. Blue Diversion: a new approach to sanitation in informal settlements. Journal of Water, Sanitation and Hygiene for Development 5, 64. doi:10.2166/washdev.2014.115

ACCEPTED MANUSCRIPT

Lempert, R.J., Collins, M.T., 2007. Managing the risk of uncertain threshold responses: comparison of robust, optimum, and precautionary approaches. Risk Anal. 27, 1009–1026. doi:10.1111/j.1539-6924.2007.00940.x

Lempert, R.J., Groves, D.G., Popper, S.W., Bankes, S.C., 2006. A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios. Management Science 52, 514–528. doi:10.1287/mnsc.1050.0472

Lempert, R.J., Popper, S.W., Bankes, S.C., 2003. Shaping the next one hundred years: new methods for quantitative, long-term policy analysis. RAND, Santa Monica, CA.

Loowatt, 2014. The Loowatt System [WWW Document]. Loowatt. URL http://loowatt.com/global-sanitation-madagascar-toilets/ (accessed 2.1.16).

Luethi, C., Morel, A., Tilley, E., Ulrich, L., 2011. Community-Led Urban environmental sanitation planning: CLUES - Complete guidelines for decision-makers with 30 tools. Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland, WSSCC, Geneva, Switzerland and UN-HABITAT, Nairobi, Kenya.

Lundgren, K., Kuklane, K., Venugopal, V., 2014. Occupational heat stress and associated productivity loss estimation using the PHS model (ISO 7933): a case study from workplaces in Chennai, India. Glob Health Action 7, 25283.

Lüthi, C., McConville, J., Kvarnström, E., 2010. Community-based approaches for addressing the urban sanitation challenges. International Journal of Urban Sustainable Development 1, 49–63. doi:10.1080/19463131003654764

Mattioli, C., 2014. Crowd Sourced Maps: Cognitive Instruments for Urban Planning and Tools to Enhance Citizens' Participation, in: Contin, A., Paolini, P., Salerno, R. (Eds.), Innovative Technologies in Urban Mapping, Sxi — Springer for Innovation. Springer, New York, pp. 145–156.

May, R.J., Maier, H.R., Dandy, G.C., Fernando, T.M.K.G., 2008. Non-linear variable selection for artificial neural networks using partial mutual information. Environmental Modelling & Software 23, 1312–1326. doi:10.1016/j.envsoft.2008.03.007

Murungi, C., van Dijk, M.P., 2014. Emptying, Transportation and Disposal of feacal sludge in informal settlements of Kampala Uganda: The economics of sanitation. Habitat International 42, 69–75. doi:10.1016/j.habitatint.2013.10.011

Oah, S., Lee, J.-H., 2011. Effects of Hourly, Low-Incentive, and High-Incentive Pay on Simulated Work Productivity: Initial Findings With a New Laboratory Method. Journal of Organizational Behavior Management 31, 21–42. doi:10.1080/01608061.2011.541820

Paterson, C., Mara, D., Curtis, T., 2007. Pro-poor sanitation technologies. Geoforum, Pro-Poor Water? The Privatisation and Global Poverty Debate 38, 901–907. doi:10.1016/j.geoforum.2006.08.006

Prüss, A., Kay, D., Fewtrell, L., Bartram, J., others, 2002. Estimating the burden of disease from water, sanitation, and hygiene at a global level. Environmental health perspectives 110, 537–542.

Ranger, N., Garbett-Shiels, S.-L., 2012. Accounting for a changing and uncertain climate in planning and policymaking today: lessons for developing countries. Climate and Development 4, 288–300. doi:10.1080/17565529.2012.732919

Rubinstein, R.Y., Kroese, D.P., 2011. Simulation and the Monte Carlo method. John Wiley & Sons.

Saeys, Y., Inza, I., Larrañaga, P., 2007. A review of feature selection techniques in. Bioinformatics 23, 2507–2517. doi:10.1093/bioinformatics/btm344

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global sensitivity analysis: the primer. John Wiley & Sons.

Sanergy, 2015. The Sanergy Model [WWW Document]. The Sanergy Model. URL http://saner.gy/ (accessed 1.2.16).

Sasmita, M., 2010. Engineering Economics and Costing. PHI Learning Pvt. Ltd.

Sharholy, M., Ahmad, K., Mahmood, G., Trivedi, R.C., 2008. Municipal solid waste management in Indian cities – A review. Waste Management 28, 459–467. doi:10.1016/j.wasman.2007.02.008

Sheppard, M., 2012. Fit all valid parametric probability distributions to data - File Exchange - MATLAB Central [WWW Document]. URL http://www.mathworks.com/matlabcentral/fileexchange/file_infos/34943-fit-all-valid-parametric-probability-distributions-to-data (accessed 10.26.14).

The MathWorks, Inc., 2014. Sequential feature selection.

Tilley, E., Strande, L., Lüthi, C., Mosler, H.-J., Udert, K.M., Gebauer, H., Hering, J.G., 2014. Looking beyond Technology: An Integrated Approach to Water, Sanitation and Hygiene in Low Income Countries. Environ. Sci. Technol. 48, 9965–9970. doi:10.1021/es501645d

Tobias, R., O'Keefe, M., Künzle, R., Gebauer, H., Gründl, H., Morgenroth, E., Pronk, W., Larsen, T.A., 2017. Early testing of new sanitation technology for urban slums: The case of the Blue Diversion Toilet. Sci. Total Environ. 576, 264–272. doi:10.1016/j.scitotenv.2016.10.057

Tumwebaze, I.K., Niwagaba, C.B., Günther, I., Mosler, H.-J., 2014. Determinants of households' cleaning intention for shared toilets: Case of 50 slums in Kampala, Uganda. Habitat International 41, 108–113. doi:10.1016/j.habitatint.2013.07.008

UNDP, 2014. Human development report 2014: sustaining human progress- reducing vulnerability and building resilience. United Nations Development Programme, New York.

UNHABITAT, 2013. Planning and Design for Sustainable Urban Mobility: Global Report on Human Settlements 2013, Global Report on Human Settlements. Routledge, Abingdon.

Urich, C., Rauch, W., 2014. Exploring critical pathways for urban water management to identify robust strategies under deep uncertainties. Water Research 66, 374–389. doi:10.1016/j.watres.2014.08.020

WHO/UNICEF, 2012. Progress on drinking water and sanitation. Joint Monitoring Programme update 2012. World Health Organization; UNICEF.

XRunner, 2014. The Xrunner system [WWW Document]. http://www.xrunner-venture.com. URL http://www.xrunner-venture.com/the-system/4583965217 (accessed 1.31.16).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Appendix 1: Full list of variables and abbreviations

Sanitation service model (see Appendix 2 for the full definition of the model)

	Unit	Description
c_C	*\$/d*	Costs, capital
$c_{C,0}$	*\$*	Initial capital costs (vehicle and equipment)
c_E	*\$/d*	Costs, energy
$cap_{Fac,max}$	*kg*	Facility maximum holding capacity
cap_{Sys}	*user*	System capacity
$cap_{V,max}$	*kg*	Vehicle maximum transport capacity
C_{tot}	*\$/d*	Costs, total
$C_{tot,user}$	*\$/user/d*	Costs per user, total
D	Facilities/d	Service demand (number of facilities requiring servicing)
$d_{Fac\text{-}Fac}$	*m*	Distance facility-facility
i_{Serv}	*d*	Mean service interval of facilities
ι_{Cap}	-	Interest rate, cost of capital
m_F	*kg(feces)/user/d*	Product accumulation rate, feces
m_{Tot}	*kg/user/d*	Product accumulation rate, total
m_U	*l(urine)/user/d*	Product accumulation rate, urine
N	Facilities/d	Total number of serviced facilities
n_{Fac}	-	Total number of facilities
n_{sim}	-	Number of simulated days
$n_{user,Fac}$	User/facility	User number per facility
N_{user}	User/d	Number of served users
n_W	-	Number of workers in a vehicle crew
p_{fuel}	*\$/l*	Fuel price
r_{fuel}	*l/hr*	Fuel consumption
s_1	*\$/facility*	Payment, service person, payment scheme 1
s_2	*\$/d*	Payment, service person, payment scheme 2
S_{Fac}	*kg*	Storage level of a facility
S_V	*kg*	Storage/Load of the vehicle
T	*h/d*	Spent working hours
t	*h*	Total (service and travel) time per facility
$t_{Fac\text{-}fac}$	*h*	Travel time between facilities
t_{max}	*h/d*	Maximum working hours
t_{Path}	*h*	Travel (walking) time on path
$t_{Project}$	*yr*	Project life span
t_{Serv}	*h*	Service time per facility
$t_{Tr\text{-}Fac}$	*h*	Travel time from treatment to facility and *vice-versa*
V_{Path}	*Km/h*	Service person's walking speed on paths
v_{Roads}	*km/h*	Vehicle speed on roads

Probabilistic spatial analysis for contextualization (see section 2.3.3 for a description of the probabilistic spatial analysis)

d_{Path}	m	Distance on foot path
$d_{Tr\text{-}Fac}$	m	Distance treatment point to facility
f_{Path}	-	Fraction of facilities on path
l_{Fac}	-	Number of candidate locations for toilet facilities
l_{Tr}	-	Number of candidate locations for treatment facilities
n_N	-	Number of neighboring facilities considered for probabilistic route analysis
P	-	Probability of finding the next facility full
P'	-	Probability of finding the next facility empty

ACCEPTED MANUSCRIPT

Abbreviations

FSA	Feature selection algorithm
PS 1	Payment scheme 1 (performance based payment)
PS 2	Payment scheme 2 (fixed daily payment)
RRP	Resource Recovery Plant
S1	Setting 1, Kampala (high housing density), Uganda
S2	Setting 2, Kampala (low housing density), Uganda
S3	Setting 3, Raipur, India
PL1	Performance level 1. Performance above the target (i.e., < 5 % failure probability and < 0.015$/user/d) on 95 % of the days
PL2	Performance level 2. Performance above the target (i.e., < 5 % failure probability and < 0.015$/user/d) on 50 % of the days

Appendix 2: Numerical sanitation service model

This section presents the numerical model that describes the servicing of $1 \dots n_{Fac}$ facilities. Facilities are serviced by a demand-driven sanitation service system. The service system consists of one or multiple workers traveling with a service vehicle between full facilities and emptying them during one or multiple daily service rounds. During a service round, j facilities are visited and emptied (until the vehicle is full) and k service rounds are performed on day i until the daily maximum working hours, $t_{max,i}$ are reached, or all full facilities are serviced. Hence the subscripts i, j indicate variables or parameters that change within the stochastic model between days (e.g. daily working hours), or from facility to facility (e.g. facility fill level, service time). This section introduces how the two key outputs of the model, system capacity and system costs are calculated

Modeling system capacity

The storage level $S_{Fac,i,j}$ of the facilities $1 \dots n_{Fac}$ on day i is generated by

$$\boldsymbol{S_{Fac,i,j} = S_{Fac,i-1,j} + m_{tot,i-1,j}} \qquad \textbf{EQ 6}$$

where $m_{tot,i-1,j}$ is the product accumulation in a facility and $S_{Fac,i-1,j}$ the storage level, both on the previous day. $m_{tot,i-1,j}$ is the accumulated mass of both, feces and urine,

$$\boldsymbol{m_{Tot,i-1,j} = \left(m_{U,i-1,j} + m_{F,i-1,j}\right) * n_{user,Fac}} \qquad \textbf{EQ 7}$$

where $n_{user,Fac}$ is the user number per toilet facility. The service demand D_i (i.e. number of facilities with depleted storage capacity) is defined at the beginning of each day. D_i considers both, facilities that reached their maximum storage capacity on day i and the full facilities that remained un-serviced on the previous day $i-1$ (D_{i-1}).

ACCEPTED MANUSCRIPT

$$\boldsymbol{D_i = \sum S_{Fac,i,j} > cap_{Fac,max} + D_{i-1}} \qquad \textbf{EQ 8}$$

where $cap_{Fac,max}$ is the maximum filling level of the storage container.

Times are calculated as follows in the model. $t_{i,j}$. the time to service a facility, is defined as

$$t_{i,j} = t_{Serv,i,j} + t_{Fac-Fac,i,j} \qquad \textbf{EQ 4}$$

where $t_{Serv,i,j}$ is the service time required to empty a facility and $t_{Fac-Fac,i,j}$ is the travel time between two facilities. If a facility is located on a footpath the service time increases to the time required to travel from the road network to the facility on a footpath:

$$\boldsymbol{t_{Serv,i,j} = t_{Serv,i,j} + t_{Path,i,j}} \qquad \textbf{EQ 5}$$

Travel times are calculated from travel distances and the velocity of the vehicle on roads

$$\boldsymbol{t_{Fac-Fac,i,j} = \frac{d_{Fac-Fac,i,j}}{v_{Roads,i}}} \qquad \textbf{EQ 6}$$

or from the velocity of the service worker on a footpath (in case facilities are located on footpaths)

$$\boldsymbol{t_{Path,i,j} = \frac{d_{Path,i,j}}{v_{Path,i}}} \qquad \textbf{EQ 7}$$

Whether a facility is on a footpath or not is determined according to a probability value that equals the measured fraction of households located on the footpath (f_{Path}).

The model equations M1-M12 shown in Table 5 explain the simulation of a service day. M defines the number of users that are served on a day *i*. From this number, also a hypothetical system capacity

ACCEPTED MANUSCRIPT

on that day can be calculated by multiplying the $N_{user,i}$ with the service interval i_{Serv}, which is 3.5 days for the proposed system, i.e., twice per week.

Table 5: Numerical sanitation service model. Variables in their order of appearance are: n_{sim} – simulation period, $n_{Fac,served,k}$ – facilities served on service round k. T_i – spent working hours, $t_{Tr-Fac,i,j}$ – travel time from treatment point to facility, $S_{V,k}$ – vehicle load, $cap_{V,max,i}$ - maximum vehicle load, $t_{max,i}$ - maximum working hours, D_i – service demand, $t_{i,j}$ – facility service time (including travel), $S_{Fac,i,j}$ – storage in facility, N_i – number of serviced facilities, $N_{user,i}$ – user served per day

Code representation	Equation number	Description
$\boldsymbol{for}\ i < n_{sim}$		*Loop through all simulated days*
$\boldsymbol{if}\ n_{Fac,served,k} = 1;$		
$T_i = t_{Tr-Fac,i,j}$	**M 1**	*Calculate travel time to 1st facility*
$\boldsymbol{while}\ S_{V,k} < cap_{V,max,i}\ (condition\ A)$ $\boldsymbol{and\ while}\ T_i < t_{max,i}\ (condition\ B)$ $\boldsymbol{and\ while}\ D_i > 0\ (condition\ C)$	**M 2**	*Continue service round if: (A) working hours, (B) vehicle capacity are not exceeded and, (C) there are full facilities left*
$T_i = T_i + t_{i,j}$	**M 3**	*Increase time*
$S_{V,k} = S_{V,k} + S_{Fac,i,j}$	**M 4**	*Add facility content to vehicle load*
$S_{Fac,i,j} = 0$	**M 5**	*Reset facility storage*
$D_i = D_i - 1\ ;\ N_i = N_i + 1$	**M 6**	*Decrease facility counter*
$\boldsymbol{if}\ S_{V,k} \geq cap_{V,max,i}$		*Vehicle capacity exceeded*
$T_i = T_i + t_{Tr-Fac,i,j}$	**M 7**	*Return to treatment point*
$\boldsymbol{elseif}\ T_i \geq t_{max,i}$		*Daily working hours exceeded*
$T_i = T_i + t_{Tr-Fac,i,j}$	**M 8**	*Return to treatment point*
$i = i + 1$	**M 9**	*Begin new day*
$\boldsymbol{elseif}\ D_i = 0$		*All facilities emptied*
$T_i = T_i + t_{Tr-Fac,i,j}$	**M 10**	*Return to treatment point*
$i = i + 1$	**M 11**	*Begin new day*
$\boldsymbol{end\ if}$		
$\boldsymbol{end\ while}$		
$N_{user,i} = N_i * n_{user,Fac}$	**M 12**	*Calculate number of users served on day i*
$\boldsymbol{end\ for}$		

Modeling system costs

The calculation of daily costs includes the calculation of fixed (e.g. equipment) and variable costs (e.g. labor, energy), divided by the number of serviced users

$$\mathbf{C_{Tot,i}} = (\mathbf{c_{W,i}} + \mathbf{c_{E,i}} + \mathbf{c_{C,i}})/\mathbf{N_{users,i}} \qquad \textbf{EQ 8}$$

Capital costs are calculated as an annuity for the equipment (Sasmita, 2010):

$$c_C = c_{c,0} * \left[\frac{(1+\iota_{Cap})^{t_{Project}} * \iota_{Cap}}{(1+\iota_{Cap})^{t_{Project}} - 1}\right] * \frac{1}{365} \qquad \textbf{EQ 9}$$

ACCEPTED MANUSCRIPT

Energy costs are a function of daily travel times [hr] and the fuel consumption rate [l/hr]

$$c_{E,i} = r_{fuel,i} * \left(\sum t_{Tr-Fac,i,j} + \sum t_{Fac-Fac,i,j}\right) * p_{fuel,i} \qquad \textbf{EQ 10}$$

We implement two payment schemes for the workers (PS 1: performance based payment. Payment is a function of the number of serviced facilities,N_i. PS 2: fixed, daily payment). Cost for paying a service person is

$$c_{W,i} = \begin{cases} s_1 * N_i * n_W \ (\text{PS 1}) \\ s_2 * n_W \ (\text{PS 2}) \end{cases} . \qquad \textbf{EQ 11}$$

- Transport-based sanitation services in informal settlements are modeled
- The framework identifies most sensitive parameters for successful implementation
- The stochastic numerical model is contextualized using satellite imagery
- Scenarios can be developed and tested for a large number of future conditions
- The framework is tested for three informal settlements in Africa and India

ACCEPTED MANUSCRIPT