Water Research 196 (2021) 116997

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Predictive models using “cheap and easy” field measurements: Can they fill a gap in planning, monitoring, and implementing fecal sludge management solutions?

Barbara J. Ward[a,b,*], Nienke Andriessen[a], James M. Tembo[c], Joel Kabika[c], Matt Grau[d], Andreas Scheidegger[a], Eberhard Morgenroth[a,b], Linda Strande[a]

[a] Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
[b] Institute of Environmental Engineering, ETH Zürich, Zürich, Switzerland
[c] Department of Civil and Environmental Engineering, School of Engineering, University of Zambia, Lusaka, Zambia
[d] Department of Physics, ETH Zürich, 8093, Zürich, Switzerland

ARTICLE INFO

Article history:
Received 10 November 2020
Revised 19 February 2021
Accepted 1 March 2021
Available online 3 March 2021

Keywords:
Random forest
machine learning
image analysis
sanitation
WASH
fecal sludge

ABSTRACT

The characteristics of fecal sludge delivered to treatment plants are highly variable. Adapting treatment process operations accordingly is challenging due to a lack of analytical capacity for characterization and monitoring at many treatment plants. Cost-efficient and simple field measurements such as photographs and probe readings could be proxies for process control parameters that normally require laboratory analysis. To investigate this, we evaluated questionnaire data, expert assessments, and simple analytical measurements for fecal sludge collected from 421 onsite containments. This data served as inputs to models of varying complexity. Random forest and linear regression models were able to predict physical-chemical characteristics including total solids (TS) and ammonium (NH_4^+-N) concentrations, and solid-liquid separation performance including settling efficiency and filtration time (R^2 from 0.51-0.66) based on image analysis of photographs (sludge color, supernatant color, and texture) and probe readings (conductivity (EC) and pH). Supernatant color was the best predictor of settling efficiency and filtration time, EC was the best predictor of NH_4^+-N, and texture was the best predictor of TS. Predictive models have the potential to be applied for real-time monitoring and process control if a database of measurements is developed and models are validated in other cities. Simple decision tree models based on the single classifier of containment type can also be used to make predictions about citywide planning, where a lower degree of accuracy is required.

1. Introduction

The sanitation needs of 1/3 of the world's population are met by non-sewered sanitation technologies, which can only provide protection of human and environmental health if the accumulated fecal sludge is adequately managed (WHO 2017, 2018). To achieve this, characterization of process control parameters at treatment plants is required to ensure safe and efficient treatment (Bassan and Robbins 2014, von Sperling et al. 2020b). This includes projections for quantities and qualities of influent when designing a treatment plant, and routine monitoring for process control and compliance of effluent standards during operation. Although there has been considerable research focused on improving characterization and monitoring for centralized, sewer-based wastewater treatment in high-income countries, this is still lacking for fecal sludge treatment (Corominas et al. 2018, Klinger et al. 2019, WHO 2017, Yoo et al. 2008). A key obstacle is the lack of accessible analytical capacity, as there are few fecal sludge treatment plants (FSTPs) with onsite laboratories, and supply chains for procurement of chemicals are often complex and unreliable (Bassan et al. 2015, Bousek et al. 2018, Klinger et al. 2019).

Solid-liquid separation is a key step in fecal sludge treatment, and FSTPs are designed for settling and dewatering which typically precede subsequent treatment of liquid and solid fractions, and is most commonly achieved by settling-thickening tanks and/or drying beds, but could also entail mechanical dewatering, geo-

* Corresponding author.
E-mail address: barbarajeanne.ward@eawag.ch (B.J. Ward).

https://doi.org/10.1016/j.watres.2021.116997

textiles, or new innovative technologies. The characteristics of fecal sludge are highly variable, resulting in correspondingly inconsistent settling and dewatering performance (Cofie et al. 2006, Gold et al. 2016). Without methods to predict the variable characteristics of sludge arriving for treatment, adjustments cannot be made to plant operations. The results are clogged drying beds or filter membranes, wasted space and decreased treatment capacity (Klinger et al. 2019). To address these problems, real-time monitoring for adaptive process control is required (Ward et al. 2021). Parameters relevant for citywide planning and optimized process control at FSTPs include physical-chemical characteristics of influent (e.g., total solids (TS), chemical oxygen demand (COD), and ammonium nitrogen (NH_4^+-N) (Bassan and Robbins 2014).

Research is being conducted into alternatives to laboratory-based analysis for fecal sludge, but these alternatives are not yet well established and little quantitative research has been done (Bousek et al. 2018). However, several interesting correlations between field measurements and laboratory-based measurements have been reported. For example, electrical conductivity (EC) and pH have been correlated with settling performance and dewatering time of fecal sludge, suggesting that these might be possible field measurements to use as indicators of expected solid-liquid separation performance (Gold et al. 2018, Junglen et al. 2020, Ward et al. 2019). Data gathered by questionnaire, including containment type (i.e., pit latrine or septic tank), presence of household water connection, source (i.e., household or commercial), and household income have been linked to physical-chemical characteristics such as TS and COD in several cities. However, these correlations are empirical and may differ between cities (Englund et al. 2020, Strande et al. 2021a, Strande et al. 2018, Tembo 2019). Expert assessments of color and odor are informally used as indicators of level of sludge stabilization, and it is generally believed that perceptible differences in color and odor are linked to different physical-chemical properties and dewatering performance (Hartenstein 1981, Schoebitz et al. 2016). However, the capacity of color and odor to monitor fecal sludge characteristics has, to our knowledge, never before been quantified and is not well documented in the literature. Quantifying relationships between possible field measurements and laboratory-based measurements is the first step in establishing alternative field-based methods, but in order get the most utility out of field-based measurements, predictive models can be employed.

Based on experience using software sensors in the wastewater treatment sector, it should be possible to develop predictive models for fecal sludge using the types of field measurements previously discussed (Dürrenmatt and Gujer 2012, Tyralis et al. 2019). Currently, operators at FSTPs do not use predictive models, but may use expert knowledge or data collected from emptiers to make decisions about operation, process control, and maintenance. For example, mixing sludge from households and public toilets in a pre-determined ratio to achieve more consistent settling behavior (Cofie et al. 2006) or varying the dose of polymer flocculant for pit latrine sludge and septic tank sludge, based on observations of the differences in their solids contents (Ward et al. 2021). Very little research has been published on predicting fecal sludge characteristics, however it has been proposed that questionnaire data can be used to model estimated loadings for planning new FSTPs (Strande et al. 2021a). A combination of data-driven and mechanistic models based on questionnaire data have demonstrated the ability to predict TS in fecal sludge with the goal of improved citywide planning (Englund et al. 2020). In order for predictive models to reduce dependence on laboratory-based characterization of fecal sludge, they must be accurate enough for routine monitoring and/or process control at FSTPs. However, so far, no such models have been reported in the literature. In this study, we use a large dataset that incorporates a combination of analytical field measurements, questionnaire data, and expert assessments to assess the suitability of predictive models for optimizing fecal sludge treatment.

The objective of this study was to investigate to what extent predictive models using field measurements can be used as a proxy for laboratory-based methods, based on a dataset of samples taken from 421 onsite containments in Lusaka, Zambia. We present results of field and laboratory data collection, and the predictive performance of the models, and discuss which models and field measurements would be appropriate for characterizing and monitoring fecal sludge in different scenarios including citywide planning, routine FSTP operations, and real-time process control of treatment technologies.

2. Materials and Methods

2.1. Sample collection

465 fecal sludge samples were collected in situ from 421 onsite containments (septic tanks and pit latrines) throughout the city of Lusaka, Zambia from September to December of 2019. Sampling replicates were taken at 8% of containments. A sampling plan was developed according to the method described in (Strande et al., 2021a). Household sampling sites were selected based on a modified version of stratified random sampling (including equal proportions of strata based on geological and demographic characteristics) in the non-sewered part of the city using ArcMap software (version 10.6). Non-household sampling sites (public toilets, offices, schools, and malls) were selected throughout the city based on local expert knowledge. Questionnaires were administered to the owner, tenant, or person in charge of operating and maintaining the system at each sampling site. Questionnaires included questions about designation of containment type, toilet type, water connection, and a number of other factors (questions available in SI). Answers were recorded with the KoBo Toolbox mobile phone app.

Because of the range of sludge consistencies present in Lusaka, different sampling devices were used to sample from pit latrines and septic tanks. A conical metal pit sampler, adapted from the design developed by James Tembo, presented in (Koottatep et al. 2021), was used to sample pit latrines. A composite sample was produced using three 1 L samples collected from the bottom (or the maximum reach of the sampler, 3 m), middle, and top of the pit. After homogenizing the composite sample in a bucket, a 0.9 L sample was taken for analysis. For septic tanks, a 3 m long core sampler adapted from the design from CDD Society, presented in Kootatep et al. (2021), was used. A composite sample was produced by emptying the contents of the core sampler into a bucket, homogenizing the contents, and taking a 0.9 L sample for analysis. Samples were transferred to a cooler for transport to the laboratory, where they were stored at 4°C until analysis. Detailed information about the sampling process, along with photographs of the sampling devices is provided in Kootatep et al. (2021).

2.2. Sample analysis

2.2.1. Sample processing

Incoming samples were homogenized thoroughly by shaking/stirring, and were divided into two portions – one to be blended before physical-chemical characterization, and the other to be analyzed for solid-liquid separation performance (avoiding blending in this case, so as to not disrupt any flocs or structure of the sludge which would influence settling and dewatering behavior).

2.2.2. Physical-chemical characterization

Samples were homogenized in a blender (3 minutes, medium setting). Foam height was measured immediately after blending using a ruler held to the wall of the blender. pH, EC, TS and volatile solids (VS) were analyzed according to the standard methods (APHA 2017). Density was measured by determining the mass of 25 mL of sample. COD was measured using the closed reflux titrimetric method (APHA 2017). Samples that could not be analyzed immediately were preserved by acidifying to pH $\leq$ 2 using H_2SO_4. NH_4^+-N was measured using the phenate method, following swirling with activated charcoal and filtration to remove filtrate color and residual turbidity (APHA 2017). Total organic carbon (TOC_{solids}) and total Kjeldahl nitrogen (TKN_{solids}) were measured on the dried solids as indicators of the potential for resource recovery of the solids, e.g., as a feedstock for composting (Al-Muyeed et al. 2017) or larvae production (Gold et al. 2020). Blended sludge was dried in a 105°C oven for 48 hours and dried solids were analyzed by an external laboratory (UNZA Department of Soil Science). TOC_{solids} was measured using the Walkely-Black procedure, with the endpoint determined by titrimetric method (Sparks et al. 2020). TKN_{solids} was measured using standard methods (Cottenie et al. 1982). A certified reference material, ISE 952 clay was used as a quality control standard.

2.2.3. Qualitative assessment of odor and color

Odor was assessed during sample processing after stirring, before blending. Samples were allowed to reach room temperature, at which point the lid of the sample container was removed, and the laboratory technician used standard chemical wafting technique (NRC 2010). Odors were classified into three categories: “fresh” (smells like fresh excreta, urine, or feces), “stabilized” (smells like compost, soil, or well biodegraded anaerobic digester sludge), or “middle” (sample odor falls somewhere between these categories). To keep the expert classifications as consistent as possible, the same laboratory technician was responsible for all odor classifications. Color was grouped into three qualitative categories: “light brown”, “black”, and “medium brown” (when sample falls somewhere between these categories). One laboratory assistant was responsible for all color classifications for consistency.

2.2.4. Solid-liquid separation performance

Supernatant turbidity was quantified as a metric of settling efficiency. Following centrifugation at 3,300 × g for 20 minutes, supernatant was decanted and turbidity was measured using a Hach 2100N turbidity meter and reported in NTU, adapted from the method described in (Mikkelsen and Keiding 2002). This was intended to represent the maximum possible reduction in suspended solids due to prolonged settling or mechanical solid-liquid separation. Capillary suction time (CST) was quantified as a metric of filtration time. CST was measured using a Triton 319 Multi-CST apparatus with 18 mm funnel, according to Method 2710 G (APHA 2017), as adapted in (Velkushanova et al. 2021a). CST values are reported in seconds, and have been standardized by subtracting the CST of deionized water. TS in the dewatered sludge cake was quantified as a metric of maximum solids content achievable by dewatering, and was defined as the dry solids content in the dewatered sludge cake after dewatering via centrifugation at 3,300 × g for 20 minutes (Gold et al. 2018). Supernatant was decanted for turbidity measurement and centrifuge tubes were left standing upside down on an absorbent cloth for 5 minutes to completely remove any free water before transferring solids out of the centrifuge tube for TS characterization.

Samples were screened for consistency, and only liquid and slurry samples (225 samples) were analyzed for solid-liquid separation performance, according to designations of sludge consistency presented in (Velkushanova and Strande 2021).

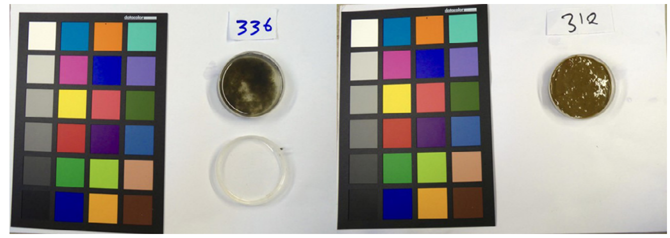


Fig. 1. Examples of setup for sludge color and texture quantification. The color checker chart shown on the left side of the photographs was used to perform standardized color correction. Color and texture information were extracted from thumbnails of the samples isolated from the color corrected images. Left: Top petri dish contains bulk sludge, bottom petri dish contains supernatant after centrifugation. Right: Petri dish contains bulk sludge.

2.2.5. Color and texture quantification

Color and texture measurements were taken by analyzing photographs of 10 mL aliquots of sludge. To our knowledge, this is a novel method, developed as an alternative to the standard colorimetric/photometric method of color detection. It is intended to be used as a field measurement. Shown in Fig. 1, 10 mL samples of unblended sludge and supernatant after centrifugation (when available) were poured into 10 mL petri dishes and placed on a white background next to a color checker chart (Datacolor Spydercheckr 24) (see Figures S1 and S2 in SI for additional photographs of the setup).

Samples were photographed in a designated location with consistent LED lighting (3 × 6W LED lamps, 3000 K (warm white color)) using a Panasonic DMC-TZ70 camera in JPEG format. JPEG format was selected over RAW in order to enable the broader application of color and texture assessment to photographs collected with smartphone cameras in the future. Photographs were post-processed for color correction using the color checker chart’s standard RGB values and assuming CIE Standard Illuminant D65 (Schanda 2007). The color correction was performed using the Python Colour library (Version 0.1.5) using the transformation presented in Cheung et al. (2004). Sludge and supernatant color swatches were manually isolated from the color corrected photographs, and the average RGB color was calculated by taking the independent averages of the R, G, and B values in the swatch. The average sludge and supernatant color values were then converted to hue, saturation, and value (HSV).

Texture analysis was performed using texture measures derived from a Grey Level Co-occurrence Matrix (GLCM) with the Python scikit-image library (version 0.17.2) (Van der Walt et al. 2014). Six common texture measures were calculated for each photograph: contrast, dissimilarity, homogeneity, correlation, angular second momentum (ASM), and energy. The mathematical descriptions of the texture measures are described in Haralick (1979) and Hall-Beyer (2017). All code related to color correction and texture analysis is provided along with the open dataset for this paper (link in SI).

2.2.6. QA/QC

According to QA/QC procedures, triplicate laboratory analyses were made for every 10th measurement for COD, NH_4^+-N, and supernatant turbidity, and every 5th measurement for TS, VS, and TS in the dewatered cake. For TOC_{solids} and TKN_{solids}, duplicate measurements were made every 7th measurement. Every CST measurement was replicated four times. Relative standard error on the replicates was (mean, 90th percentile): COD (8%, 16%), NH_4^+-N (4%, 9%), TS (5%, 15%), VS (4%, 8%), TOC_{solids} (8%, 13%), TKN_{solids} (18%, 34%), CST (5%, 11%), turbidity (4%, 9%), and TS in the dewatered cake (10%, 24%).

Table 1
Cost-efficient and simple field measurements that were evaluated as predictors for the more costly laboratory-based analytical methods presented in Table 2.

Field measurements (model inputs)	Details
Questionnaire	Categorical, collected via questionnaire
containment type	septic tank, pit latrine
toilet type	wet flush, dry toilet
water connection on premises	yes, no
source	household, non-household
Expert assessments	Categorical, collected by laboratory assistant
odor	stabilized, middle, fresh
color (qualitative)	black, medium brown, light brown
Simple analytical	Continuous, collected by the following equipment (units):
EC	probe (mS/cm)
pH	probe
foam height	ruler (mm)
color (quantitative)	camera (hue, saturation, value)
texture	camera (contrast, dissimilarity, homogeneity, correlation, angular second moment (ASM), energy)
supernatant color	camera (hue, saturation, value)

Table 2
Laboratory-based measurements.

Laboratory-based measurements (target parameters)	Details
Solid-liquid separation performance:	
Supernatant turbidity following centrifugation	Measure of settling efficiency (NTU)
CST	Measure of filtration time (seconds)
TS in dewatered cake following centrifugation	Measure of maximum water removal (% dry solids)
Physical-chemical characteristics:	
Complete sample	
COD	(g/L)
NH_4^+-N	(g/L)
TS	(% dry solids)
VS	(% of total solids)
Dried solids	
TOC_{solids}	(% of total solids)
TKN_{solids}	(% of total solids)

2.3. Data interpretation

Median values within a category were considered to be significantly different from one another if the 95% confidence intervals around the medians did not overlap (Chambers et al. 1983). Confidence intervals (CI) were calculated using the formula

$$CI = median \pm 1.57 \times \frac{IQR}{\sqrt{n}}$$

Where IQR is the interquartile range and n is the number of data points in the category (Chambers et al. 1983). Confidence intervals around the median are represented as notches on boxplots (Figures S3 and S4 in SI).

2.4. Predictive models

Cost-efficient and simple to execute field measurements were evaluated in this study to determine whether they could be used to predict more costly laboratory-based analytical measurements. Field measurements evaluated as inputs to predictive models are detailed in Table 1. Questionnaire data is categorical and quantitative. Expert assessments are categorical, qualitative, and were assigned by trained laboratory technicians. Simple analytical measurements are quantitative, continuous-valued measurements collected using a probe, a ruler, or a camera. Although simple analytical measurements were performed in a laboratory as part of this study, they are all able to be measured in the field without laboratory equipment, and thus are labeled field measurements. The laboratory-based measurements that are desired outputs of the predictive models (target parameters) are outlined in Table 2.

Data analysis, modelling, and visualization was performed using Python 3 (Van Rossum and Drake 2009). Models were implemented in scikit-learn (version 0.23.2) and statsmodels (version 0.12.1) packages in Python 3. Models, along with the complete dataset are available open source with this publication (link in SI).

Three types of models were evaluated in this study: i) a simple decision tree model, ii) a linear regression model, and iii) a random forest model. These models were chosen as they represent a range of complexities and unique advantages and disadvantages. Simple decision tree models are based on the decision tree presented in Strande et al. (2021a), which uses the median value of the target parameter in a category (e.g. median TS in septic tanks) to predict future values in that category. Simple decision tree models are reflective of the way that operational decisions may currently be made at treatment plants (Cofie et al. 2006, Ward et al. 2021). They have the advantage of easy interpretation and visualization, but predictive capabilities are often poor. Linear regression models can also be easily interpreted, and are good descriptions of systems with linear behavior. In contrast, random forest models are a widely applied black-box machine learning algorithm that can deal with non-linearities and interactions, but cannot be interpreted directly (Hastie et al. 2009). A thorough description of the specifics for each model type is provided in the SI.

Model evaluation took part in two steps: a) the identification of relevant field measurement inputs for predicting each target laboratory-based measurement using a reduced dataset, and b) the final assessment of the performance of the three model types for each target, using expanded datasets.

Model inputs were evaluated using a reduced dataset (n = 244) comprising all inputs included in Table 1, with samples with missing input data removed. Supernatant color was included as an input for models predicting solid-liquid separation performance, but not for models predicting physical-chemical characteristics, as supernatant color data was collected only for the subset of samples that were evaluated for solid-liquid separation performance.

For each target parameter, models were generated for every possible combination of maximal four inputs, and the performance of each model was evaluated using cross-validated R^2 and root mean squared error (RMSE) (5-fold cross validation, repeated 20 times). The input combination with the highest cross-validated R^2 was considered the best. Preference was given to models with fewer inputs, so if inputs could be removed from the model without a relevant decrease in R^2 (at two decimal places), those inputs were not included in the best model. Relative importance of inputs was evaluated by comparing the R^2 of models built with and without the input. The relative strengths of the inputs included in the best models were evaluated by comparing the R^2 of single-input models. The input with the largest R^2 was labeled the 'strongest predictor' if the R^2 of that model was at least 75% of the R^2 of the best multi-input model. Supporting predictors were defined as inputs that are included in the best model and increase the model R^2 when included as model inputs along with the strongest predictor. Detailed information about model performance and input importance are included in the SI (Tables S3-S9).

Input selection was also dictated by model type. The simple decision tree model (as defined in Strande et al. (2021a)) was designed to use only categorical data, so only questionnaire data and expert assessments were used as inputs to this model. The linear regression model was evaluated for all inputs (questionnaire data, expert assessments, and simple analytical measurements), except the texture parameters dissimilarity and ASM. These were removed after pre-screening for collinearity with other inputs (Pearson coefficient > 0.85), as they were strongly correlated with other texture parameters. The random forest model was evaluated for all inputs.

After the relevant inputs for predicting each target had been identified, final model performance was determined for each target. The dataset for the final evaluation included only the relevant inputs that were used in the best models (of each model type) for predicting a specific target, in order to maximize the number of data points used to train and evaluate each model. This allows the performance of the three model types in predicting a specific target to be compared. Cross-validated R^2 and RMSE were used in the final performance evaluation.

3. Results & Discussion

Data collected with cost-efficient and simple to execute field measurements were used to develop predictive models of the more expensive, laboratory-based analysis (i.e. physical-chemical characteristics and solid-liquid separation performance), in order to evaluate whether the less expensive methods could be used as proxies or partial replacements in data collection. The results and discussion are presented in the order of: 1) presentation of the collected field- and laboratory-based data; 2) a comparison of the developed models and best predictors; and 3) implications for use in characterization of fecal sludge for different fecal sludge management scenarios, including citywide planning, routine FSTP operations, and real-time process control of treatment technologies.

3.1. Overview of characteristics and trends in collected data

Results of the characterization of fecal sludge from septic tanks and pit latrines are presented in Tables 3 and 4. Overall, the values were highly variable, and did not follow a normal distribution, which is consistent with other studies (Strande et al. 2021b). Median and mean field measurements, physical-chemical characteristics and solid-liquid separation performance metrics measured in this study are within the expected range based on reported median and mean values in the literature for sludge from septic tanks and pit latrines in Lusaka (Tembo 2019, Tembo et al. 2019), Kampala, Uganda

Table 3
Results of simple analytical field measurements grouped by septic tank and pit latrine. Literature values for comparison are ranges of reported mean or median values.

	Field measurements														
	EC	pH	foam height	color			supernatant color			texture					
	(mS/cm)		(mm)	H	S	V	H	S	V	cont.	dissim.	homog.	ASM	energy	corr.
SEPTIC TANKS															
mean	2.6	7.64	5.62	53	28	21	54	17	87	2.69	0.44	0.85	0.35	0.56	0.63
std	2.5	0.45	6.05	15	19	20	24	15	10	7.87	0.55	0.09	0.20	0.17	0.22
median	**1.8**	**7.66**	**4.00**	**50**	**23**	**13**	**51**	**11**	**87**	**0.35**	**0.28**	**0.86**	**0.32**	**0.57**	**0.66**
25%	1.4	7.49	0.00	45	14	10	49	7	83	0.25	0.20	0.83	0.20	0.45	0.45
75%	2.7	7.83	10.00	60	39	19	56	23	93	0.77	0.36	0.91	0.44	0.66	0.81
n	202	202	202	200	200	200	181	181	181	197	197	197	197	197	197
literature values:															
reported means	2.3-15.4[a,b,c,d]	6.9-7.9[a,b,c,d,e]													
reported medians	2.0-13.5[a,c,d]	7.4-7.8[a,c,d]													
PIT LATRINES															
mean	14.2	7.59	2.71	50	46	21	41	69	66	9.33	1.10	0.74	0.23	0.45	0.50
std	5.3	0.53	6.63	9	20	9	10	18	16	9.50	0.73	0.11	0.16	0.15	0.15
median	**14.5**	**7.73**	**0.00**	**48**	**50**	**20**	**42**	**74**	**68**	**6.36**	**1.02**	**0.72**	**0.18**	**0.42**	**0.49**
25%	11.2	7.39	0.00	45	30	15	35	67	56	1.50	0.41	0.66	0.12	0.34	0.39
75%	17.2	7.96	2.00	54	61	26	48	82	78	14.81	1.61	0.82	0.29	0.54	0.59
n	207	207	207	203	203	203	46	46	46	198	198	198	198	198	198
literature values:															
reported means	12.1-14.6[a]	7.1-8.2[a,d,g]													
reported medians	12.0-13.6[a]	7.1-8.2[a,d,g]													

(a) Gold et al. 2018, (b) Gold et al. 2016, (c) Ward et al. 2019, (d) Englund et al. 2020, (e) Heinss et al. 1999, (f) Bassan et al. 2013, (g) Tembo 2019, (h) Semiyaga et al. 2017, (i) Tembo et al. 2019

Table 4
Results of laboratory-based analysis grouped by septic tank and pit latrine. Literature values for comparison are ranges of reported mean or median values.

	Laboratory-based measurements								
	Supernatant turbidity (NTU)	CST (s)	TS in dewatered cake (% ds)	COD (g/L)	NH4-N (g/L)	TS (% ds)	VS (% of TS)	TOC_{solids} (% of TS)	TKN_{solids} (% of TS)
SEPTIC TANKS									
mean	180	86	22.3	72.1	0.5	4.8	53.5	10.9	2.2
std	230	132	22.2	56.9	0.7	6.6	20.2	2.7	1.2
median	**100**	**42**	**14.8**	**53.3**	**0.3**	**2.0**	**51.8**	**11.0**	**2.1**
25%	40	13	7.4	32.0	0.1	0.6	41.8	9.1	1.4
75%	240	114	26.6	93.9	0.5	6.8	64.9	13.0	3.0
n	179	172	157	165	202	189	181	142	142
literature values:									
reported means		97.6[c]	11-18[a,c]	7.6-43.0[a,b,d,e,f]	0.18-0.6[a,b,d]	0.8-7.2[a,b,c,d,e,f]	48.3-76[a,b,c,d,e,f]		
reported medians		57.8[c]	12-14[a,c]	7.5-35[a,d]	0.24-0.63[a,d]	0.6-2.6[a,c,d]	52.7-75.5[a,c,d]		
PIT LATRINES									
mean	850	707	22.0	122.6	3.0	14.7	56.4	10.9	2.3
std	800	788	16.6	65.5	1.5	8.2	16.9	3.1	1.2
median	**650**	**468**	**18.2**	**121.1**	**3.0**	**14.8**	**59.0**	**9.9**	**2.2**
25%	390	304	9.4	82.1	2.1	9.8	44.3	9.1	1.4
75%	1000	834	31.8	156.5	3.7	18.9	69.2	13.6	2.8
n	46	45	43	154	206	195	193	124	124
literature values:									
reported means		179-1485[c,h]	18-31.8[a,c,h]	10.9-129[a,d,f,g]	1.4-3.2[a,d]	0.9-19[a,d,f,g,i]	43.2-63[a,d,f,g]		
reported medians		179[c]	17-30[a,c]	9.8-127.2[a,d,g]	1.3-3.1[a,d]	1.1-17[a,d,g]	52-64[a,d,g]		

(a) Gold et al. 2018, (b) Gold et al. 2016, (c) Ward et al. 2019, (d) Englund et al. 2020, (e) Heinss et al. 1999, (f) Bassan et al. 2013, (g) Tembo 2019, (h) Semiyaga et al. 2017, (i) Tembo et al. 2019

(Englund et al. 2020, Gold et al. 2018, Semiyaga et al. 2017), Dakar, Senegal (Gold et al. 2016, Ward et al. 2019), Dar es Salaam, Tanzania (Ward et al. 2019), Durban, South Africa (Radford et al. 2015), Ouagadougou, Burkina Faso (Bassan et al. 2013), Accra, Ghana (Heinss et al. 1999), Bangkok, Thailand (Heinss et al. 1999), Manila, Philippines (Heinss et al. 1999), and Hanoi, Vietnam (Englund et al. 2020, Gold et al. 2018). The complete dataset for this study is available open source with this publication: https://doi.org/10.25678/00037X .

As a first step in developing the predictive models, correlations and trends within the data were investigated. Values for EC, foam height, color (saturation), supernatant color (saturation), texture (contrast, dissimilarity, homogeneity, ASM, energy, and correlation), supernatant turbidity, CST, COD, NH_4^+-N, TS, and TOC_{solids} were significantly different for pit latrine and septic tank sludge, whereas pH, TKN_{solids}, and TS in the dewatered cake were not significantly different (based on 95% confidence intervals around the medians).

Dependencies between questionnaire categories of containment type, toilet type, water connection, and source exist. Illustrated in Fig. 2, there is a dependency between containment type (i.e. pit latrine or septic tank) and toilet type (i.e. dry or wet), as water is used to convey excreta from toilets to septic tanks, whereas water is not required for conveyance to pit latrines. In a similar fashion, sites that had water connections were also more likely to have septic tanks. 55% of households had pit latrines and 45% septic tanks, whereas non-household establishments had a majority of septic tanks (80%). It was surprising that 2.5% of septic tanks were associated with dry toilets, and 12% with no water connection. This calls into question what people actually mean when they report 'septic tank' and brings attention to the apparent disparity between common assumed definitions of septic tanks and how they are actually defined in the field. In the future, instead of septic tank and pit latrine, descriptors of the actual containment technology (for example: lined, unlined, baffled, presence of overflow) are likely to provide more accurate and globally comparable descriptions (Johnston and Slaymaker 2020).

Questionnaire results for categories of containment type, toilet type, water connection, and source were related to laboratory-based measurements. Pit latrines, dry toilets, and sites with no water connection yielded correspondingly less diluted sludge (significantly higher median COD, NH_4^+-N, and TS) with poorer settling and filtration performance (higher supernatant turbidity and CST). Septic tanks, wet toilets (either pour-flush or cistern flush), and sites with a water connection yielded correspondingly more diluted sludge (significantly lower median COD, NH_4^+-N, and TS) with better settling and filtration performance (lower supernatant turbidity and CST). Because of the strong correlations between containment type and toilet type, only containment type is further included as a potential model input. Sludge from households was less diluted (higher median COD, NH_4^+-N, and TS), and had poorer settling and filtration performance (higher supernatant turbidity and CST) compared with sludge from non-household sources (Figures S2 and S3). These results contrast with observations in earlier studies in Dakar and Dar es Salaam, where sludge from public toilets (non-household) had poorer settling and dewatering performance than sludge from households (Ward et al. 2019), but agree with observations in Kampala that there was a difference between physical-chemical properties (COD, TS) of sludge from household and non-household sources (Strande et al. 2018). This suggests that source (i.e. household or non-household) may sometimes be a useful predictor of sludge characteristics, but the predictive relationships are different in different cities.

The expert assessment categories of color and odor show dependencies with each other, and with questionnaire categories. 98% of 'fresh' smelling samples were light or medium brown, while 74%

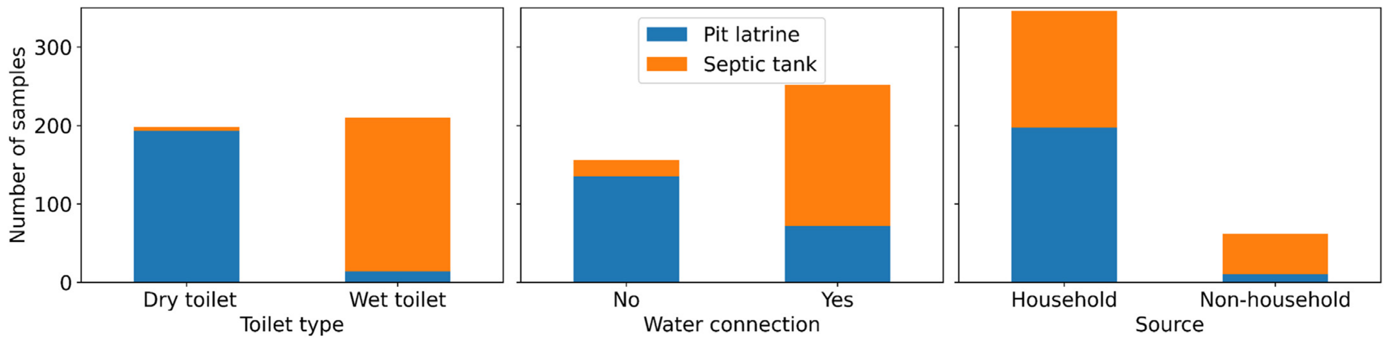


Fig. 2. a) Distribution of pit latrines and septic tanks connected to dry and wet toilets, b) distribution of pit latrines and septic tanks at locations with and without onsite water connections, and c) distribution of pit latrines and septic tanks at households and non-household sites.

of 'stabilized' smelling samples were black (see Fig. S5). The majority of pit latrine sludge had a 'fresh' or 'middle' odor (79%) and was medium or light brown (74%), while the majority of septic tank sludge had a 'stabilized' odor (75%) and was black (74%) (see Fig. S6).

The expert assessment categories color and odor were associated with laboratory-based measurements. There were significant differences in the median supernatant turbidity, CST, NH_4^+-N, TS, COD, VS, and TOC_{solids} between 'fresh'/light brown and 'stabilized'/black samples (Figures S2 and S3). Interestingly, our results support common practitioner wisdom that has, to our knowledge, not been quantified in the literature: fecal sludge with higher organic matter content (COD, TOC, VS) was associated by odor and color as "fresh" sludge, whereas sludge labeled "stabilized" had correspondingly significantly lower organic matter. 'Fresh' sludge also took significantly longer to dewater and had poorer settling performance. This corroborates practitioner observations from the field that fresh sludge dewaters and settles more poorly than more stabilized sludge (Cofie et al. 2006, Heinss et al. 1999, Ward et al. 2019). This finding is interesting, because it offers insight into the transformation of organic matter in sludge as it stabilizes. The specific shift from the "fresh" feces odor to the more stabilized "barnyard" or "manure" odor during stabilization is a result of bacterial metabolism of organic acids, which produce the smells associated with fresh feces, leaving behind phenolic and sulfur-containing compounds, which are associated with the odor of stabilized feces (Lin et al. 2013, Starkenmann 2017). These results can direct future research into the transformation of organic matter in the sludge during stabilization and its effect on fundamental mechanisms controlling dewatering.

Relationships to simple analytical measurements are discussed in detail in Section 3.2.2.

3.2. Comparison of model input combinations and performance

The collected data were evaluated to see if field measurements (Table 1) could be used to predict laboratory-based measurements (Table 2). A range of field measurements were evaluated as inputs in the predictive models, including questionnaire data such as containment type (i.e. pit latrine or septic tank), source (i.e. household or non-household)), expert assessments such as odor and color (qualitative), and simple analytical measurements such as EC, pH, foam height, color (quantitative), texture, and supernatant color. Table 5 summarizes the combinations of field measurements that were selected as relevant inputs in every model type for each target parameter.

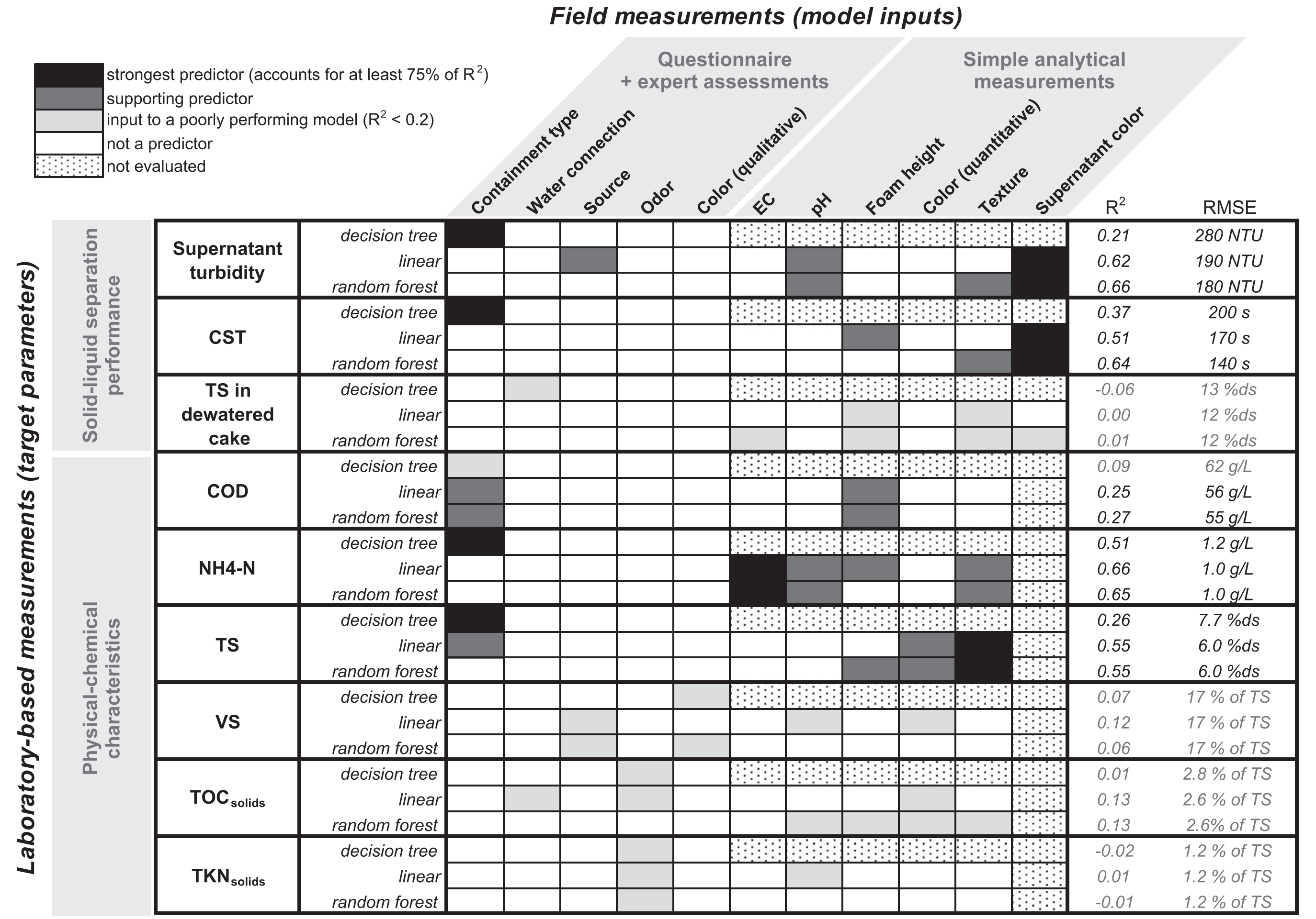
The strongest predictor for each model is shown in black, and supporting predictors, which increase the model performance when included as model inputs, are shown in grey. Model inputs included in poorly performing models ($R^2 < 0.2$) are shown in light grey. The model performance metrics on the right of the figure are cross-validated R^2 and RMSE for models using the highlighted field measurements as inputs. Since simple decision tree models are based on single inputs, the input associated with the highest performing model for each target was designated as the strongest predictor. Tables showing model prediction accuracy and error depending on the inclusion of different predictors are included in the SI (Tables S3-S9).

3.2.1. Model performance

As seen in Table 5, in predicting a given target, simple decision tree models did not perform as well, while linear regression models and random forest models generally improved model fit and reduced error. Decision tree models were able to account for 21-51% of the variance in supernatant turbidity, CST, NH_4^+-N, and TS. Linear models improved on the predictions made by decision tree models for supernatant turbidity, CST, NH_4^+-N, TS, and COD, with 195%, 38%, 178%, 29%, and 112% increases in model fit (R^2) and 32%, 15%, 10%, 17%, and 22% reductions in prediction error (RMSE) respectively. Random forest models improved on the predictions made by decision tree models for supernatant turbidity, CST, NH_4^+-N, TS, and COD, with 214%, 73%, 200%, 27%, and 112% increases in model fit (R^2), respectively, and 36%, 30%, 11%, 17%, and 22% reductions in prediction error (RMSE), respectively. Random forest models outperformed linear models only in predicting the solid-liquid separation performance metrics supernatant turbidity and CST. Random forest models improved on the predictions made by linear models for supernatant turbidity and CST with 6% and 25% increases in R^2 and 5% and 18% reductions in RMSE, respectively. It was unexpected that random forest models did not provide higher accuracy predictions than linear models for most parameters, as random forest models are able to capture nonlinear interactions between predictors (Hastie et al. 2009), and nonlinear relationships have been observed in models of nutrients in wastewater effluent (Castrillo and García 2020). It is possible that given access to a larger training dataset, random forest model performance could exceed that of linear models.

Something to consider when selecting a predictive model is the trade-off between increased predictive power and interpretability of the model. Simple decision tree models and linear regression models tend to be less flexible and provide less robust predictions for complex nonlinear datasets, but it is relatively easy to understand the relationships between a field measurement and the target parameter. In contrast, nonlinear methods such as random forest models are more flexible and may produce better predictions, but at the expense of model interpretability (Hastie et al. 2009). In cases where linear models and random forest models provide comparable predictions, linear models may be preferred for their relative simplicity. In cases where data will be collected routinely, and a dataset will continue to grow in size and complexity, random

Table 5
Field measurements identified as relevant model inputs to predict laboratory-based measurements, and corresponding R^2 and RMSE of models run with selected inputs. Strongest predictors (accounting for at least 75% of R^2) shown in black, supporting predictors shown in grey, inputs to poorly performing models (R^2<0.2) shown in light grey, inputs not used in model shown in white, and inputs not evaluated shown with grey dots. Model performance metrics, R^2 and RMSE for each target parameter by type of model are displayed on the right.



Group	Target parameter	Model	R^2	RMSE
Solid-liquid separation performance	Supernatant turbidity	decision tree	0.21	280 NTU
		linear	0.62	190 NTU
		random forest	0.66	180 NTU
	CST	decision tree	0.37	200 s
		linear	0.51	170 s
		random forest	0.64	140 s
	TS in dewatered cake	decision tree	-0.06	13 %ds
		linear	0.00	12 %ds
		random forest	0.01	12 %ds
Physical-chemical characteristics	COD	decision tree	0.09	62 g/L
		linear	0.25	56 g/L
		random forest	0.27	55 g/L
	NH4-N	decision tree	0.51	1.2 g/L
		linear	0.66	1.0 g/L
		random forest	0.65	1.0 g/L
	TS	decision tree	0.26	7.7 %ds
		linear	0.55	6.0 %ds
		random forest	0.55	6.0 %ds
	VS	decision tree	0.07	17 % of TS
		linear	0.12	17 % of TS
		random forest	0.06	17 % of TS
	TOC_{solids}	decision tree	0.01	2.8 % of TS
		linear	0.13	2.6 % of TS
		random forest	0.13	2.6% of TS
	TKN_{solids}	decision tree	-0.02	1.2 % of TS
		linear	0.01	1.2 % of TS
		random forest	-0.01	1.2 % of TS

forest models may be preferred for their ability to extract nonlinear relationships from larger datasets.

3.2.2. Suitability of field predictors

Although questionnaire data and expert assessments were useful inputs for the decision tree models, they were only incrementally helpful to include in the linear and random forest models. Containment type was the strongest predictor for decision tree models, explaining 21-51% of the variance in supernatant turbidity, CST, NH_4^+-N, and TS. For all linear and random forest regression models with R^2 higher than 0.5, a single field measurement (the strongest predictor) was responsible for at least 75% of the prediction accuracy (black in Table 5). Although the inclusion of questionnaire data (containment type and source) did contribute to increased fit for some linear and random forest models, the simple analytical parameters were always the strongest predictors for these types of models. This indicates that the predictive detail represented by the separation into septic tank and pit latrine or household and non-household is largely captured by the differences in the physical-chemical compositions of sludge from septic tanks and pit latrines or household and non-household sources. The simple analytical field measurements may be better model inputs than questionnaire data or expert assessments because they are continuous instead of categorical and are thus able to provide higher resolution information.

Supernatant turbidity and CST were predicted primarily by supernatant color. Supernatant color contributed 86% and 87% of the linear and random forest model fits, respectively, in predictions of supernatant turbidity, and 75% and 90% of linear and random forest model fits, respectively, for predictions of CST. Including texture as an input further improved random forest model predictions of both supernatant turbidity and CST. Our results support and quantify previous observations of a relationship between qualitatively measured supernatant color and settling and filtration performance in fecal sludge from Dakar and Dar es Salaam (Ward et al. 2019). This relationship is hypothesized to be a result of high concentrations of suspended and soluble organic matter, which also contribute to high supernatant turbidity and filter clogging during dewatering (Ward et al. 2019). Field measurements can be selected to suit available technical and financial resources. Although supernatant color was the strongest predictor of supernatant turbidity and CST, it may not be an ideal field predictor in every case, as it requires settling prior to measurement. However, settling tests are also simple analytical tests that can be performed in the field with low-cost equipment in less than an hour (e.g. Imhoff cones or settling columns) (Junglen et al. 2020). It is not surprising that supernatant turbidity is strongly related to supernatant color. Now that this relationship has been quantified for fecal sludge in Lusaka, it seems promising for estimations of effluent turbidity using color in photographs as a proxy, and could replace turbidity measurements where spectrophotometers are not available.

NH_4^+-N was predicted primarily by EC, which contributed 91% and 82% of the linear and random forest model fits, respectively. (The linear model is simple: EC (mS/cm) * 0.2 = NH_4^+-N (g/L), see Fig. S7). The inclusion of texture and pH in the models further improved predictions. There is precedent in the literature for the use of EC as a predictor of ammonia in manure and fecal sludge. Onsite measurement of EC has been suggested as a good proxy for predicting nutrient concentrations, including ammonia nitrogen ($R^2 = 0.91$) in swine manure (Suresh et al. 2009). For fecal sludge from pit latrines and septic tanks in Uganda, Vietnam, and Japan, Gold et al. (2018) observed a linear correlation ($R^2 = 0.6$) between NH_4^+-N and EC. Based on these results, onsite measurement of EC using a conductivity probe is a promising option for cheaper and less resource-intensive monitoring of NH_4^+-N.

TS was predicted primarily by texture, which contributed 81% and 86% of the linear and random forest model fits, respectively. The inclusion of color (quantitative) further improved both model types, and including containment type additionally improved linear model predictions. These results make sense based on scientific knowledge that sludge surfaces gain texture as they dry, changing from a smooth liquid to a lumpy slurry, to a rough semi-solid or solid. The physical transformation of sludge as it dries following dewatering has been well characterized for sediment sludges on drying beds, and TS has been shown to be predictable to a high level of accuracy ($R^2 > 0.92$) using machine learning based on texture analysis (Bodun et al. 2000). Our results indicate that using texture measurements extracted from photographs for prediction of TS could be useful in eliminating the time- and equipment-intensive laboratory analysis of TS for fecal sludge.

The target parameters VS, TOC_{solids}, TKN_{solids}, and TS in the dewatered cake were not able to be predicted using any of the models. These parameters, along with COD (for which models were only able to account for ~25% of the variation in the data), are all associated with the sludge organic matter. It appears that the field parameters evaluated in this study are not sufficient to fully characterize and predict the organic components in fecal sludge. This is consistent with our previous observations that it was difficult to predict variation in dewatered cake solids and VS in fecal sludge from Dakar and Dar es Salaam (Ward et al. 2019). This was hypothesized to be partly due to the influence of soil and solid waste on measured VS and sludge dewatering behavior. A rapid field measurement to predict silica content as a proxy for soil could be a possible solution. (Miller et al. 2012) proposed the use of field portable infrared spectrometers to predict silica content in coal dust. Another possibility could be monitoring organic matter composition using in situ fluorescent sensors. This has been demonstrated in the field, where COD and soluble COD in the effluent of decentralized wastewater treatment systems were able to be monitored using fluorescence as a proxy measurement (Mladenov et al. 2018). Fluorescent sensing may also be able to provide better predictions of the composition and transformation of organic matter in sludge during stabilization; fluorescent peaks have been associated with the concentrations of high molecular weight humic substances and other organic matter, and thus may be a promising method for monitoring level of stabilization (Mladenov et al. 2018, Yao et al. 2016). Further fundamental understanding of the mechanisms controlling fecal sludge stabilization and dewatering will likely be instructional in identifying possible field predictors.

Models can also be adapted to incorporate new field measurements, especially in cases when practitioners have identified them to be operationally relevant. For example, we found that including density as an input can significantly improve the prediction accuracy of the TS model (random forest model prediction improves to $R^2 = 0.70$, RMSE = 4.8 %ds). For practitioners who have access to a penetrometer for field density measurements as described in Radford and Sugden (2014), this could be a valuable field predictor to incorporate.

3.3. Implications for the field

In this study, we evaluated three types of models with increasing levels of complexity. The selection of the best model for different applications will depend on the required level of accuracy and on the available resources. Here, we explore where and how each model type could be relevant, considering several situations where predictive models could be useful for characterization and monitoring, as summarized in Fig. 3.

Scenario 1: Community to citywide planning

Information provided by simple decision tree models may be adequate for the level of detail required for community to citywide planning. Incorporating questionnaire data such as containment type into simple decision tree models can already provide a moderate amount of information about sludge characteristics and solid-liquid separation performance. Siting decisions and designs for new FSTPs are often based on citywide averages of loadings and volumes (Klinger et al. 2019), but it has recently been suggested that incorporating decision tree models to estimate quantities and qualities of fecal sludge accumulating on a neighborhood level would greatly improve projections of loadings for planning of FSTPs, transfer stations, and regularly scheduled empting programs (Strande et al. 2021a). Our study confirms that such an approach would provide an improvement over citywide averages: a simple decision tree model based on containment type accounts for 26% of the variance in TS. Because this application does not require a high-resolution prediction, decision tree models providing low-accuracy predictions (R^2 of 0.26 for TS) maybe be adequate, as they are easy to understand conceptually and are in many cases based on information that city planners might already have access to. This can include information collected in a shit flow diagram (SFD), for example, prevalence of pit latrines and septic tanks by neighborhood, location of water taps, or data on housing density or residential/commercial land use (Peal et al. 2020).

Scenario 2: Periodic process control of fecal sludge treatment technologies

Operating a fecal sludge treatment plant requires more refined predictions than citywide planning. Combining cost-efficient and simple field measurements with linear or machine learning models contributes a further improvement in prediction accuracy, providing moderate-accuracy predictions of TS, NH_4^+-N, and settling and filtration performance (R^2 of 0.51-0.66). Operators routinely make decisions about how much sludge to load on drying beds, and when to remove dewatered/dried sludge. Improvements in consistency of solids loading and shorter, more consistent residence times could substantially increase treatment plant capacity and performance (Klinger et al. 2019, Seck et al. 2015). Currently, the drying beds at most FSTPs are operated at constant hydraulic loadings and residence times, with no monitoring of influent sludge characteristics (Klinger et al. 2019). As a result of not being able to monitor influent sludge, common operational problems due to the variability in influent TS arise. These problems include overloaded drying beds clogging and dewatering too slowly, and underloaded beds wasting space and decreasing treatment capacity (Klinger et al. 2019). Being able to adjust the loading of drying beds based on TS concentrations and dewatering time, together with monitoring of TS on the drying beds, would increase treatment performance and treatment plant capacity. This could be done by incorporation of linear or machine learning models into a smartphone app, so that pictures taken with the smartphone would provide estimates of TS concentrations and dewatering time. A more

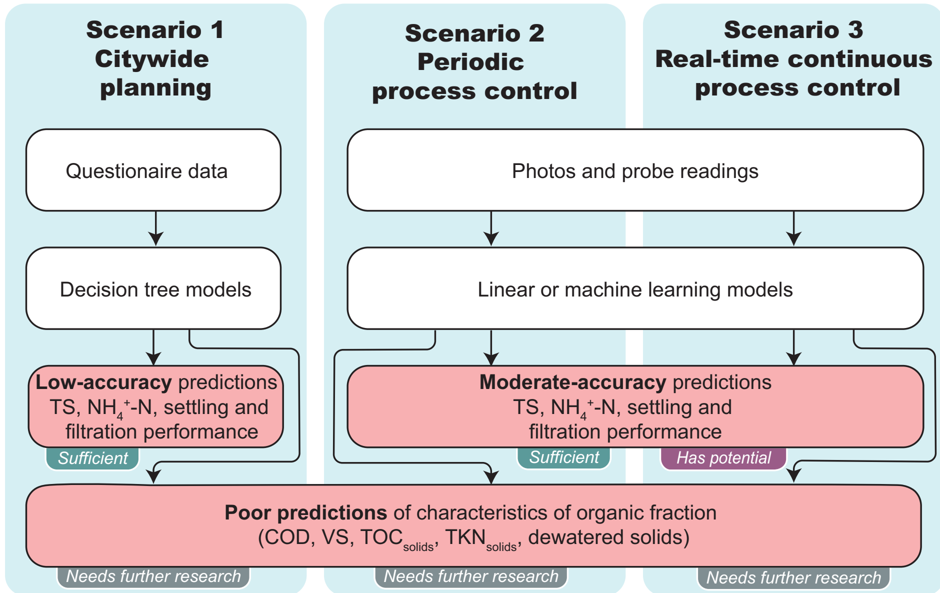


Fig. 3. This flowchart summarizes three types of planning and operation scenarios where cost efficient and simple field measurements and models could be employed. Appropriate inputs and model types for each scenario are shown as white boxes, and prediction accuracies of different outputs are shown as pink boxes. The tabs at the base of each pink box indicate whether the prediction accuracy is likely to be sufficient, has potential with further data collection, or requires additional research prior to implementation.

simple application could be employed with printed cards with example colors and textures for comparison and decision-making (von Sperling et al. 2020a).

Scenario 3: Real-time continuous process control of fecal sludge treatment technologies

The capacity to make rapid predictions based on photos and probe measurements could be a serious game-changer for processes requiring real-time monitoring. Online conditioner dosing for advanced settling and dewatering is one example (Ward et al. 2021). The current state of the art in FSM is to adjust polymer dosing flowrates based on containment type: pit latrine sludge is dosed at a higher flowrate than septic tank sludge to account for the differences in average TS of sludge in each category (Ward et al. 2021). This method provides insufficient resolution for predictions, and it is very difficult to avoid over- and under-dosing conditioner due to the high variability in sludge characteristics. In this case, a smartphone app with photo and probe inputs (obtained at the treatment facility) could be used to monitor TS. If the TS value of the influent sludge is 2% ds, the random forest model should predict within a range of 0.7 – 3.3 % ds. This may be sufficiently high resolution, depending on the propensity of the selected conditioner to overdosing. For example, using a 2000 L mixing tank at a transfer station (Rhodes-Dicker et al. 2020) and a target polymer conditioner dose of 2 mL/g TS, an actual dose between 0.7 and 3.3 mL/g TS could be achieved (see SI for calculation). This is comparable to the observed optimal dosing window for the conditioner (1-3 mL/g TS). These models are a proof of concept; they will also need to be refined depending on the technology and validated prior to use in other cities. Currently, there is a significant lack of information on fecal sludge characterization to drive the development of predictive models. The development of a global database of fecal sludge field measurements uploaded together with laboratory results using standardized methods (Velkushanova et al. 2021b) will allow for the continuous improvement of models and global applications of these predictions. Such an approach has significant potential to provide reliable characterization data and enable real-time monitoring and process control for fecal sludge treatment in cities all over the world.

4. Conclusions

Based on the findings in this study, the key conclusions are:

- Cost-efficient and simple field measurements from photos (color, texture) and probes (EC, pH) can be used as predictors of fecal sludge characteristics and solid-liquid separation performance (TS, NH_4^+-N, settling efficiency, and filtration time) when combined with linear regression and machine learning models.
- Containment type is a good predictor of fecal sludge characteristics and solid-liquid separation performance (TS, NH_4^+-N, settling efficiency and filtration time) and can be especially helpful for making low-resolution predictions when combined with simple decision tree models, e.g. for projecting loadings for FSTP design.
- Laboratory-based measurements associated with the organic matter in the sludge solids (COD, VS, TOC_{solids}, TKN_{solids}, TS in the dewatered cake) could not be predicted using the methods we evaluated. A better understanding of the organic matter in fecal sludge and how it relates to solid-liquid separation performance is needed to identify good predictors.
- Based on this proof of concept, which indicates that predictions of characteristics using photographs and probe measurements are possible, focus should be placed on validating this approach in other cities. The collection of worldwide datasets would allow for global implementation and continuously improving machine learning models. Our ongoing research includes development of an app for field practitioners that can predict fecal sludge characteristics based on pictures taken with a smartphone.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are appreciative of the team members and collaborators in Lusaka, whose effort, time, and expertise were instrumental in collecting the data in this study. Lab team: Chipo Chuunya, Samuel Mangalashi, Gift Sikwale, Bukata Chanda, Sikazwe Mutakwa, Beene Chiyumbabeenzu, Grace Chilufya, Richard Matanda. Supervising lab technicians: Dan Mkandawire, Enock Musonda Mutati, Derrick Muwowo. Dept. of soil science, UNZA: Gideon Musukwa. Head of Department CEE, UNZA: Balimu Mwiya. Sampling team: Cornelius Makai, Charles Machipisha, Enock Mumba, Kennedy Mwitumwa, Geoffrey Luanja, Chama Chabala. UNZA drivers: Eric Musonda, Betrine Chibomba, Pythias Mwelwa, Eustace Chewe. Lusaka City Council field visit representatives: Chuma Banda, Lilian Wamunyima, Juanita Mumba. Collaborators at GIZ: Chaiwe Mushauko-Sanderse, Kapanda Kapanda, Mwaba Kapema, Trevor Surridge.

Samples and questionnaire data were collected as part of the "Faecal sludge quantities and qualities (Q&Q) in Lusaka" project, supported by the German Corporation for International Cooperation GmbH (GIZ). Additional financial support was provided by Eawag, the Swiss Agency for Development and Cooperation (SDC), the Swiss National Science Foundation (SNF), and the Engineering for Development (E4D) Doctoral Scholarship Program of ETH Zurich, funded through the Sawiris Foundation for Social Development.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2021.116997.

References

Al-Muyeed, A., Oko-Williams, A., Islam, K., Ali, L., Sanyal, P.R., 2017. Co-composting of faecal sludge with solid waste to improve FSM practice in Sakhipur municipality Loughborogh, UK.
APHA, 2017. Standard Methods for the Examination of Water and Wastewater. American Public Health Association (APHA), American Water Works Association (AWWA), 23rd Edition Water Environment Federation (WEF), Washington, D.C, USA.
Bassan, M., Koné, D., Mbéguéré, M., Holliger, C., Strande, L., 2015. Success and failure assessment methodology for wastewater and faecal sludge treatment projects in low-income countries. Journal of Environmental Planning and Management 58 (10), 1690–1710.
Bassan, M. and Robbins, D.M. (2014) Faecal Sludge Management: Systems Approach for Implementation and Operation. Strande, L., Ronteltap, M. and Brdjanovic, D. (Eds.), pp. 231-253, IWA Publishing.
Bassan, M., Tchonda, T., Yiougo, L., Zoellig, H., Mahamane, I., Mbéguéré, M., Strande, L., 2013. Characterization of faecal sludge during dry and rainy seasons in Ouagadougou. Burkina Faso, Nakuru, Kenya.
Bodun, P., Shibusawa, S., Sasao, A., Sakai, K., Nonaka, H., 2000. Dredged sludge moisture prediction by textural analysis of the surface image. Journal of terramechanics 37 (1), 3–20.
Bousek, J., Skodak, M., Bäuerl, M., Ecker, G., Spit, J., Hayes, A., Fuchs, W., 2018. Development of a Field Laboratory for Monitoring of Fecal-Sludge Treatment Plants. Water 10 (9), 1153.
Castrillo, M., García, Á.L., 2020. Estimation of high frequency nutrient concentrations from water quality surrogates using machine learning methods. Water Research 172, 115490.
Chambers, J.M., Cleveland, W.S., Kleiner, B., Tukey, P.A., 1983. Graphical methods for data analysis. Wadsworth & Brooks/Cole.
Cheung, V., Westland, S., Connah, D., Ripamonti, C., 2004. A comparative study of the characterisation of colour cameras by means of neural networks and polynomial transforms. Coloration technology 120 (1), 19–25.
Cofie, O.O., Agbottah, S., Strauss, M., Esseku, H., Montangero, A., Awuah, E., Kone, D, 2006. Solid–liquid separation of faecal sludge using drying beds in Ghana: Implications for nutrient recycling in urban agriculture. Water Research 40 (1), 75–82.
Corominas, L., Garrido-Baserba, M., Villez, K., Olsson, G., Cortés, U., Poch, M., 2018. Transforming data into knowledge for improved wastewater treatment operation: A critical review of techniques. Environmental Modelling & Software 106, 89–103.
Cottenie, A., Verloo, M., Kiekens, L., Velghe, G., Camerlynck, R., 1982. Chemical analysis of plants and soils. Lab. Agroch. State Univ. Gent, Belgium.
Dürrenmatt, D.J., Gujer, W., 2012. Data-driven modeling approaches to support wastewater treatment plant operation. Environmental Modelling & Software 30, 47–56.
Englund, M., Carbajal, J.P., Ferré, A., Bassan, M., Vu, A.T.H., Nguyen, V.-A., Strande, L, 2020. Modelling quantities and qualities (Q&Q) of faecal sludge in Hanoi, Vietnam and Kampala, Uganda for improved management solutions. Journal of Environmental Management 261, 110202.
Gold, M., Dayer, P., Faye, M.C.A.S., Clair, G., Seck, A., Niang, S., Morgenroth, E., Strande, L, 2016. Locally produced natural conditioners for dewatering of faecal sludge. Environmental technology 37 (21), 2802–2814.
Gold, M., Egger, J., Scheidegger, A., Zurbrügg, C., Bruno, D., Bonelli, M., Tettamanti, G., Casartelli, M., Schmitt, E., Kerkaert, B., Smet, J.D., Campenhout, L.V., Mathys, A., 2020. Estimating black soldier fly larvae biowaste conversion performance by simulation of midgut digestion. Waste Management 112, 40–51.
Gold, M., Harada, H., Therrien, J.-D., Nishida, T., Cunningham, M., Semiyaga, S., Fujii, S., Dorea, C., Nguyen, V.-A., Strande, L., 2018. Cross-country analysis of faecal sludge dewatering. Environmental technology 39 (23), 3077–3087.
Hall-Beyer, M., 2017. GLCM Texture: A Tutorial v. 3.0 University of Calgary Calgary, Canada.
Haralick, R.M., 1979. Statistical and structural approaches to texture. Proceedings of the IEEE 67 (5), 786–804.
Hartenstein, R., 1981. Sludge decomposition and stabilization. Science 212 (4496), 743–749.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.
Heinss, U., Larmie, S.A. and Strauss, M. (1999) Characteristics of faecal sludges and their solids-liquid separation, EAWAG/SANDEC, Duebendorf, Switzerland.
Johnston, R., Slaymaker, T., 2020. Monitoring Safely Managed On Site Sanitation (M-SMOSS).
Junglen, K., Rhodes-Dicker, L., Ward, B.J., Gitau, E., Mwalugongo, W., Stradley, L., Thomas, E., 2020. Characterization and prediction of fecal sludge parameters and settling behavior in informal settlements in Nairobi. Kenya. Sustainability 12 (21).
Klinger, M., Gueye, A., Sherpa, A., Strande, L., 2019. Scoping Study: Faecal Sludge Treatment Plants in South-Asia and sub-Saharan Africa [version 1; not peer reviewed]. Gates Open Research 3.
Koottatep, T., Ferré, A., Chapagain, S., Fakkaew, K. and Strande, L. (2021) Methods for faecal sludge analysis. Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D. and Buckley, C. (Eds.), pp. 55-84, IWA Publishing.
Lin, J., Aoll, J., Niclass, Y., Velazco, M.I.s., Wünsche, L., Pika, J., Starkenmann, C, 2013. Qualitative and quantitative analysis of volatile constituents from latrines. Environmental science & technology 47 (14), 7876–7882.
Mikkelsen, L.H., Keiding, K., 2002. Physico-chemical characteristics of full scale sewage sludges with implications to dewatering. Water Research 36 (10), 2451–2462.
Miller, A.L., Drake, P.L., Murphy, N.C., Noll, J.D., Volkwein, J.C., 2012. Evaluating portable infrared spectrometers for measuring the silica content of coal dust. Journal of Environmental Monitoring 14 (1), 48–55.
Mladenov, N., Bigelow, A., Pietruschka, B., Palomo, M., Buckley, C., 2018. Using submersible fluorescence sensors to track the removal of organic matter in decentralized wastewater treatment systems (DEWATS) in real time. Water Science and Technology 77 (3), 819–828.
NRC (2010) Chemical Laboratory Safety and Security: A Guide to prudent chemical management, National Academies Press.
Peal, A., Evans, B., Ahilan, S., Ban, R., Blackett, I., Hawkins, P., Schoebitz, L., Scott, R., Sleigh, A., Strande, L., 2020. Estimating safely managed sanitation in urban areas; lessons learned from a global implementation of excreta-flow diagrams. Frontiers in Environmental Science 8.
Radford, J., Sugden, S., 2014. Measurement of faecal sludge in-situ shear strength and density. Water SA 40 (1), 183–188.
Radford, J., Underdown, C., Velkushanova, K., Byrne, A., Smith, D., Fenner, R., Pietrovito, J., Whitesell, A., 2015. Faecal sludge simulants to aid the development of desludging technologies. Journal of Water, Sanitation and Hygiene for Development 5 (3), 456–464.
Rhodes-Dicker, L., Ward, B.J., Mwalugongo, W., Stradley, L., 2020. Permeable membrane dewatering of faecal sludge from pit latrines at a transfer station in Nairobi. Kenya Journal of Environmental Management 1–28.
Schanda, J., 2007. Colorimetry: understanding the CIE system. John Wiley & Sons.
Schoebitz, L., Bischoff, F., Ddiba, D., Okello, F., Nakazibwe, R., Niwagaba, C., Lohri, C., Strande, L., 2016. Results of faecal sludge analyses in Kampala, Uganda: Pictures, characteristics and qualitative observations for 76 samples Eawag. Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland.
Seck, A., Gold, M., Niang, S., Mbéguéré, M., Diop, C., Strande, L., 2015. Faecal sludge drying beds: increasing drying rates for fuel resource recovery in Sub-Saharan Africa. Journal of Water, Sanitation and Hygiene for Development 5 (1), 72–80.
Semiyaga, S., Okure, M., Niwagaba, C., Nyenje, P., Kansiime, F., 2017. Dewaterability of faecal sludge and its implications on faecal sludge management in urban slums. International Journal of Environmental Science and Technology 14 (1), 151–164.
Sparks, D.L., Page, A., Helmke, P., Loeppert, R.H., 2020. Methods of soil analysis, part 3: Chemical methods. John Wiley & Sons.
Starkenmann, C., 2017. Springer Handbook of Odor. Springer, pp. 121–122.
Strande, L., Englund, M., Andriessen, N., Carbajal, J.P. and Scheidegger, A. (2021a) Methods for faecal sludge analysis. Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D. and Buckley, C. (Eds.), pp. 115-144, IWA Publishing.
Strande, L., Schoebitz, L., Bischoff, F., Ddiba, D., Okello, F., Englund, M., Ward, B.J.,

Niwagaba, C.B., 2018. Methods to reliably estimate faecal sludge quantities and qualities for the design of treatment technologies and management solutions. Journal of Environmental Management 223, 898–907.

Strande, L., Velkushanova, K. and Brdjanovic, D. (2021b) Methods for faecal sludge analysis. Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D. and Buckley, C. (Eds.), pp. 1-14, IWA Publishing, London, UK.

Suresh, A., Choi, H., Oh, D., Moon, O., 2009. Prediction of the nutrients value and biochemical characteristics of swine slurry by measurement of EC–Electrical conductivity. Bioresource technology 100 (20), 4683–4689.

Tembo, J., 2019. Faecal sludge characterisation for enhanced sanitation provision in peri-urban areas of Lusaka University of Zambia.

Tembo, J.M., Matanda, R., Banda, I.N., Mwanaumo, E., Nyirenda, E., Mambwe, M., Nyambe, I.A., 2019. Pit latrine faecal sludge solid waste quantification and characterization to inform the design of treatment facilities in peri-urban areas: A case study of Kanyama. African Journal of Environmental Science and Technology 13 (7), 260–272.

Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water 11 (5), 910.

Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., 2014. scikit-image: image processing in Python. PeerJ 2, e453.

Van Rossum, G., Drake, F.L., 2009. Introduction To Python 3: Python Documentation Manual Part 1. CreateSpace.

Velkushanova, K., Reddy, M., Zikalala, T., Gumbi, B., Archer, C., Ward, B.J., Andriessen, N., Sam, S. and Strande, L. (2021a) Methods for faecal sludge analysis. Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D. and Buckley, C. (Eds.), IWA Publishing, London, UK.

Velkushanova, K. and Strande, L. (2021) Methods for faecal sludge analysis. Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D. and Buckley, C. (Eds.), pp. 15-54, IWA Publishing.

Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D., Buckley, C., 2021b. Methods for faecal sludge analysis. IWA Publishing, London, UK.

von Sperling, M., Lima, E.M.M.N., de Andrade Moraes, M.A, 2020a. A simple field essay for detecting departures from expected performance in small-scale, remote or rural wastewater treatment plants. Water Science and Technology 82 (7), 1380–1392.

von Sperling, M., Verbyla, M.E., Oliveira, S.M., 2020b. Assessment of Treatment Plant Performance and Water Quality Data: A Guide for Students, Researchers and Practitioners. IWA Publishing.

Ward, B.J., Septien, S., Ronteltap, M. and Strande, L. (2021) Methods for faecal sludge analysis. Velkushanova, K., Strande, L., Ronteltap, M., Koottatep, T., Brdjanovic, D. and Buckley, C. (Eds.), pp. 85-114, IWA Publishing.

Ward, B.J., Traber, J., Gueye, A., Diop, B., Morgenroth, E., Strande, L., 2019. Evaluation of conceptual model and predictors of faecal sludge dewatering performance in Senegal and Tanzania. Water Research 167, 115101.

WHO, 2017. Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines. World Health Organization, Geneva.

WHO, 2018. Guidelines on sanitation and health. World Health Organization, Geneva.

Yao, Y., Li, Y.-z., Guo, X.-j., Huang, T., Gao, P.-p., Zhang, Y., Yuan, F., 2016. Changes and characteristics of dissolved organic matter in a constructed wetland system using fluorescence spectroscopy. Environmental Science and Pollution Research 23 (12), 12237–12245.

Yoo, C.K., Villez, K., Van Hulle, S.W., Vanrolleghem, P.A., 2008. Enhanced process monitoring for wastewater treatment systems. Environmetrics: The official journal of the International Environmetrics Society 19 (6), 602–617.