Statistical modelling of Fat, Oil and Grease (FOG) deposits in wastewater pump sumps

Nieuwenhuis, Eva; Post, Johan; Duinmeijer, Alex; Langeveld, Jeroen; Clemens, François

DOI
10.1016/j.watres.2018.02.026

Publication date
2018

Document Version
Final published version

Published in
Water Research

Water Research 135 (2018) 155–167

Contents lists available at ScienceDirect

Water Research

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

Statistical modelling of Fat, Oil and Grease (FOG) deposits in wastewater pump sumps

Eva Nieuwenhuis [a, *], Johan Post [b], Alex Duinmeijer [a, c], Jeroen Langeveld [a, b], François Clemens [a, d]

[a] Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Water Management, Section Sanitary Engineering, P.O. Box 5048 2600 GA, The Netherlands
[b] Partners4UrbanWater, Javastraat 104A, 6524, Nijmegen, MJ, The Netherlands
[c] Engineering Consultancy Municipality Rotterdam, P.O. Box 6567, 3002, Rotterdam, AN The Netherlands
[d] Deltares, Department of Industrial Hydrodynamics, P.O. Box 177, 2600, Delft, MH, The Netherlands

ARTICLE INFO

Article history:
Received 9 June 2017
Received in revised form
8 February 2018
Accepted 12 February 2018
Available online 13 February 2018

Keywords:
Sewer system
Pumping station
Fat Oil and Grease (FOG) deposits
Generalized linear mixed modelling

ABSTRACT

The accumulation of FOG (Fat, Oil and Grease) deposits in sewer pumping stations results in an increase in maintenance costs, malfunctioning of pumps and, a potential increase of wastewater spills in receiving open water bodies.

It is thought that a variety of parameters (e.g. geometry of the pump sump, pump operation, socio-economic parameters of the catchment) influences the built-up of FOG. Based on a database containing data of 126 pumping stations located in five Dutch municipalities a statistical model was built. It is shown that 3 parameters are most significant in explaining the occurrence of FOG deposits: mean income of the population in a catchment, the amount of energy (kinetic and potential) per m^3 per day and the density of restaurants, bars and hotels in a catchment. Further it is shown that there are significant differences between municipalities that can be traced back to the local 'design paradigm'. For example, in Amsterdam, the design philosophy of discharging in the pump sump under the water surface (and hence maintaining a low level of turbulence in the pump sump) results in an increase of the probability of the formation of FOG.

1. Introduction

Sewer systems are vital for public health and city life. Sewer blockages are found to be the dominant failure mechanism in sewer systems (Arthur et al., 2009; Ashley et al., 2004). In the United States, almost half of all sewer blockages are related to Fat, Oil and Grease (FOG) deposits (EPA, 2004). FOG deposits are accumulated suspended solids in sewer systems and have an adhesive character. They can become firmly attached to interior sewer pipe walls, thereby substantially reducing and sometimes even completely blocking the wastewater flow (Desilva et al., 2011). They have a grainy, sandstone-like texture with high yield strengths (Keener et al., 2008) that require intensive cleaning activities such as hydraulic jetting (Dirksen et al., 2012; Mattsson et al., 2014).

It is often thought that FOG deposits in public sewer lines result from solidified cooking oils as they are poured down the drain and cool down in downstream sewer lines. The formation mechanisms, however, have appeared to be much more complex. Keener et al. (2008) showed they are basically metallic soaps, mainly consisting of (saturated) fatty acids and calcium. Later research described the mechanism of FOG deposit formation in sewer pipes as the saponification process between calcium and free fatty acids, together with the aggregation of excess calcium, unreacted fatty acids and debris in wastewater that are drawn towards the solid core matrix of saponified solids (He et al., 2013). In addition, recent work of Gross et al. (2017) showed that FOG deposits can also be the result of acids crystallization, implying that FOG deposits can also be formed without the presence of metals.

Collected samples from different locations within the sewer network showed a wide range in physical and chemical properties (He et al., 2011; Keener et al., 2008; Nieuwenhuis et al., 2017; Shin et al., 2015; Williams et al., 2012). Different formation processes and accumulation mechanisms were suggested for different network locations (He et al., 2011; Williams et al., 2012), which is

* Corresponding author.
E-mail address: e.m.nieuwenhuis@tudelft.nl (E. Nieuwenhuis).

https://doi.org/10.1016/j.watres.2018.02.026

also in line with the recent laboratory study of Gross et al. (2017).

Both at upstream and downstream locations FOG deposits are known to accumulate:

In sewer pipes, FOG deposits typically tend to accumulate slightly above the low-flow water mark (Keener et al., 2008; Williams et al., 2012). Dirksen et al. (2012) and Dominic et al. (2013) identified sagging sewers in particular to be vulnerable to the accumulation of FOG. For lateral house connections, Post et al. (2016) showed that accumulation of FOG is the main failure mechanism.

In inverted siphons, declining parts with the presence of air pockets are particularly prone to FOG blockages (de Groot, 2015).

In pumping stations, FOG deposits with three different appearances were found. Franke et al. (2011) mentioned the floating layers of FOG, accumulating on the walls of pump sumps. These layers potentially interfere with the functioning of level regulators in the pump sump, depending on the type of level regulators. Additionally, Williams et al. (2012) collected FOG samples in the shape of 'fat balls' from the water surface of pumping stations, and Dirksen et al. (2012) mentioned the occasional detachment of bar-shaped deposits in sewer pipes. These may end up in pump sumps, as such bar-shaped deposits were observed in pump sumps during the site visits conducted for this study.

Previous studies mainly focused on the chemical aspects of FOG deposit formation. Although these studies have revealed general information on locations in public sewer systems that are prone to FOG accumulation, they did not focus on the particular sewer contexts of FOG accumulations. In addition, previous studies hardly elaborated on the probable impact of domestic disposal patterns on FOG deposits. They focused mainly on Food Service Establishments (FSEs) (Dominic et al., 2013; He et al., 2011; Williams et al., 2012), and the fishing and meat industries (Cammarota and Freire, 2006; Mattsson et al., 2014) as the main contributors to FOG problems. A recent case study in The Netherlands showed, however, that lateral house connections are more susceptible to blockages than main sewers, and that FOG is the dominant failure mechanism in lateral house connections (Post et al., 2016). This demonstrates that domestic disposal patterns are an important contributor to FOG deposits. Similarly, Wallace et al. (2017) mentioned the contribution of domestic wastewater to FOG blockages and a survey done among 127 Norwegian and Swedish sewer operators reported that respectively 48 and 22% experienced FOG-related problems in residential areas (Mattsson et al., 2014). They explicitly mentioned the severity of FOG accumulation in areas with high-rise apartment buildings and a relatively high number of immigrants (Mattsson et al., 2014).

Considering that aspects of lifestyle may be attributed to demographic groups, it is hypothesized that FOG problems are related to demographics and vary considerably in severity among catchments and in corresponding pumping stations. In addition, it is expected that pumping stations with structural configurations that enable low flow velocities are more prone to FOG build-up. The research presented here aims at finding evidence for both hypotheses. To this end, a statistical study on 126 wastewater pumping stations in five municipalities has been performed.

2. Materials

For investigating the impact of domestic disposal patterns, FOG deposits were considered on the scale of catchment areas. This allowed using demographics of catchments for studying the influence of population disposal patterns statistically.

Data on catchments and corresponding pumping stations were collected in five relatively large Dutch municipalities. Table 1 provides an overview of the participating municipalities and their general characteristics; they varied in demographics, type of catchments and pumping stations.

The dataset of residential catchments was composed in close collaboration with the municipalities, resulting in binomial data on FOG accumulation in pump sumps. Each observation is represented by one catchment and its pumping station, describing the presence or absence of severe accumulation of FOG as judged by the sewer manager. This judgement represents the state of FOG accumulation over multiple years and at least one year. It was based on a combination of 1) visual inspection by operators during regular maintenance and 2) available information about cleaning efforts required. As the municipalities did not systematically record FOG accumulation, this was the best available data.

To avoid discrepancies between cities, parameter definitions were discussed beforehand. Pumping stations without consensus on the severity of FOG accumulation or that were lacking crucial information (e.g. construction drawings) were excluded from the dataset.

2.1. Parameter selection

The investigated parameters represent general system characteristics and socio-demographic (from here on called 'demographic') characteristics that are potential indicators for FOG disposal patterns or the FOG accumulation process.

Statistical analyses require comparable parameters and one representative value per observation. The three parameters, 'vertical velocity', 'pump-on-time' and 'kinetic energy density', are therefore introduced, representing the geometry of the pumping stations and the hydraulic loading (Table 2). These parameters are related to the motion of water, and hence, are suspected to affect the accumulation of FOG.

2.1.1. Vertical velocity

The average vertical velocity v_{vert}, [mh^{-1}] is calculated as:

$$v_{vert} = \frac{Q_{pump}}{A_{sump}} \quad (1)$$

Where Q_{pump} is the pump capacity under dry weather conditions in [m^3h^{-1}], and A_{sump} [m^2] the surface area of the pump sump. Since pumping stations operate under dry weather conditions for about 80% of the time (Tukker et al., 2012), the Dry Weather Flow (DWF) is taken as the representative hydraulic loading. For variable frequency drive (VFD) pumps, the operating schemes have been provided by the municipalities, allowing to determine representative values for Q_{pump} during DWF.

2.1.2. Daily operation time

The average daily operation time, $t_{operation}$, in hours per day is calculated as:

$$t_{operation} = \frac{Q_{dwf}t}{Q_{pump}} \quad (2)$$

Where Q_{dwf} is the hourly DWF [m^3h^{-1}], t is the time [h], in this case 24 h, and Q_{pump} is the pump capacity during DWF [m^3h^{-1}].

2.1.3. Kinetic energy density

The values for kinetic energy density, i.e. the incoming energy per pump sump per day, $E_{daypump}$, in [$Jm^{-3}d^{-1}$], are based on the values for hourly DWF as provided by the municipalities. For each pumping station, hourly values for the kinetic energy, $E_{kin,h}$ [Jh^{-1}], are summed over the day and divided by the representative water volume in the pump sump, V_{sump} [m^3]:

Table 1
General characteristics of participating municipalities. The data is revealed from municipal sewer system management plans (Municipality of Almere, 2011; Municipality of Arnhem, 2009; Municipality of Rotterdam, 2011; Municipality of The Hague, 2010; Waterboard Amstel, 2010). Pumping stations were count as such when their corresponding (sub)-catchments were identified on the GIS data delivered by the municipalities.

Municipality	Number of inhabitants [-]	Pumping stations under control of municipality [-]	Length of DWF gravity sewers [km]		
			Total	Combined	Separated
Amsterdam	767,500	437	1358	525	833
Rotterdam	593,000	536	3311	1809	1502
The Hague	484,000	72	1091	845	246
Almere	188,000	178	595	0	595
Arnhem	147,000	22	464	178	286

Table 2
Selected system characteristic parameters.

Parameter	Unit	Description
City	[-]	The city where the pumping station is located
Sewer system type	[-]	The type of sewer system
Gutters	[-]	The presence of gutters arranged in a zigzag
Vertical velocity	[mh^{-1}]	The average vertical velocity in the pump sump following from to the pumping capacity under dry weather conditions
Daily operation time	[hd^{-1}]	The average operation time per day, based on DWF
Kinetic energy density	[$Jm^{-3}d^{-1}$]	Total incoming kinetic energy per unit of volume per day

$$E_{daypump} = \sum_{t=1}^{24} \frac{E_{kin,h}}{V_{sump}} \tag{3}$$

The amount of kinetic energy that got into the water in the pump sump, is calculated as the kinetic energy at the invert level of the inflowing pipe(s), $E_{kin,inv}$ [J], and the potential energy, E_{pot} [J], of the inflowing water with respect to representative water depth in the pump sump; the water level in the pump sump is assumed to be constant.

$$E_{kin} = E_{kin,inv} + E_{pot} \tag{4}$$

where E_{pot} is:

$$E_{pot} = mgh \tag{5}$$

where m [kg] is the mass of the incoming water, g is the gravitational acceleration [ms^{-2}], and h [m] the fall height of the incoming water, assuming a constant water level in the pump sump.

And where $E_{kin,inv}$ is.

$$E_{kin,inv} = \frac{1}{2}mv^2 \tag{6}$$

where m [kg] is the mass of the incoming water, and v [ms^{-1}] the flow velocity.

The velocities are derived from hourly values for the DWF, according to the hourly distribution percentages, and the cross-sectional area of flow:

$$v = \frac{Q_{dwf}}{A} \tag{7}$$

where Q_{dwf} [m^3h^{-1}] is the hourly DWF, and A [m^2] is the cross-sectional area of flow. The velocity, v [mh^{-1}], is assumed to be constant for every hour and, and the incoming DWF is assumed to be equally divided among all inlet pipes.

The cross-sectional area, A [m^2], depends on the water depth at the location of the inlet during the particular hour. This is derived from the representative water depth in the pump sump, z [m] (i.e. the water depth following from the water level in between the switch-on and switch-off levels of the DWF pump), the invert level of the inlet pipe, z_i [m], and the average water depth in the pipe at the location of inflow during the particular hour, d [m]. For the calculations of the cross-sectional area and/or the flow velocity and corresponding kinetic energy, three situations for representative water depths, z, at the location of inlet are distinguished, see Fig. 1.

For, $z \leq z_i$ the outlet of the pipe is classified as 'free outflow'. Close to the end of such pipes, flow conditions are critical, implying that the non-dimensional Froude number, F_r, is known and specified as:

$$F_r = \frac{v_c}{\sqrt{g \cdot d_m}} = 1 \tag{8}$$

where v_c [ms^{-1}], is the critical flow velocity, g [ms^{-2}] is the gravitational acceleration, and d_m [m] is the hydraulic mean depth, specified as the cross-sectional area of flow per flow width at the water surface. For such flow conditions, the empirical equation of Straub (1978) applies (9) and the critical depth d_c [m] is derived:

$$\frac{d_c}{D} = 0.567 \cdot \frac{Q_{dwf}^{0.506}}{D^{1.264}} \tag{9}$$

where Q_{dwf} [m^3s^{-1}] is the hourly DWF, D [m] is the diameter and d_c [m] the critical depth, where $0.02 < \frac{d_c}{D} \leq 0.85$.

Thereafter, using geometric and trigonometric equations, the hydraulic mean depth, d_m [m], as displayed in Fig. 1 is determined, and from (8), the critical flow velocity v_c [ms^{-1}] is derived.

For one pumping station the value is slightly below the lower limit $\frac{d_c}{D} = 0.007$, and for seven pumping stations this value is exceeding the upper limit $\frac{d_c}{D} \leq 1.31$. In these cases, $\frac{d_c}{D}$ is assumed to be equal to the lower and upper limits, as the specified conditions are only violated for minimum and maximum DWF values. The possible influence of the tail water is neglected, and the water depth at the outflow is assumed to be equal to the critical depth, thus neglecting the drawdown effect.

For $z < z_i < z_i + D$, water depth d [m] is used from (10), with the value of d [m] confined by d_c [m].

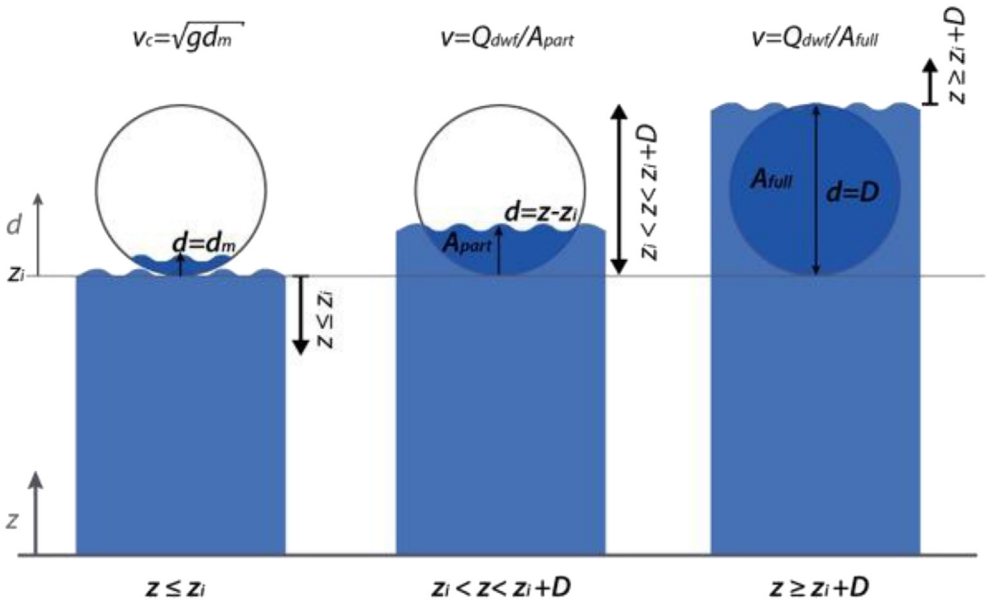

Fig. 1. Different representative water depth scenarios for calculating the kinetic energy. The figure shows the front view of the inlets.

$$d = z - z_i \tag{10}$$

For $z \geq z_i + D$, full pipe flow is considered.

Geometrical details, like the shape of the pump sump or the position of inflow are not considered, as the nature of statistical analysis does not allow for such details.

2.1.4. Demographic data

Online available geographical data from Statistics Netherlands on neighbourhood level was used to obtain weighted demographic data per catchment. The geographical maps were composed from data from the Key Registers Cadaster and regional data from Statistics Netherlands (Statistics Netherlands and Kadaster, 2012). Data from the year 2012 was used, as this was the most recent dataset covering all parameters needed. Merging data from different years was infeasible, due to changes over the years in the borders of administrative neighbourhoods.

Table 3 provides an overview of the potential explanatory parameters selected from the Statistics Netherlands' database.

Calculations are performed with Quantum GIS software, version 2.0.1-Dufour (QGIS, 2013). Using the Geoprocessing Intersect tool, a GIS layer with the (contributing areas of) neighbourhoods in each individual catchment was created.

Further data processing is performed using R statistics software, version 0.99.893 (R Core Team, 2016). The database shows missing values; data points that were identified as 'nihil' were replaced by zero, and data points that were identified as 'susceptible to reliability and secrecy' were replaced by 'NA' (not available).

The total number of inhabitants for each neighbourhood was calculated, based on the population density and the surface area of each neighbourhood, as derived from QGIS calculations.

Representative values for the number of Food Service Establishments (FSE) were derived by summing the 'average number of restaurants, cafes and cafeterias within a travel distance of 1 km for the inhabitants'. This value was divided by the surface area, to obtain a representative value for the FSE density, The household density for each catchment was calculated by taking the number of households divided by the surface area of each neighbourhood.

Estimations of the demographic characteristics per catchment were obtained by weighing the characteristics according to the catchment's population that the contributing neighbourhoods contained. Using the catchment weights, characteristics per catchment were derived in proportion to their populations. The

Table 3
Selected demographic parameters; derived from geographical data on neighbourhood level (Statistics Netherlands and Kadaster, 2012).

Parameter	Unit	Description
Population density	[km^{-2}]	Population per unit of area
Household density	[km^{-2}]	Total number of households per unit of area
Household size	[-]	Average number of total inhabitants per household
Non-western immigrants	[%]	The percentage of immigrants with non-western origin
Rental properties	[%]	Percentage of rental properties
Housing association properties	[%]	Percentage of rental properties owned by housing associations
Personal income (based on total population)	[·1000€]	Average personal income per person based on total population
Personal income (based on working population)	[·1000€]	Average personal income per person based on people with an annual income
Low income population	[%]	Percentage of households belonging to the group with the 40% lowest disposable incomes
High income population	[%]	Percentage of households belonging to the group with the 20% highest disposable incomes
Below social minimum	[%]	Percentage of households belonging to the group that has an income that is below the social minimum as established in the political decision-making
FSE density	[km^{-2}]	Average number of restaurants, cafes and cafeterias within a travel distance of 1 km for each inhabitant

numbers of inhabitants per catchment were based on the population density of the neighbourhoods and the (contributing) surface areas of the neighbourhoods. After pre-processing of the data, the database of the selected pumping and their system characteristics was merged with the demographic catchment data.

2.2. Resulting dataset

Table 4 provides an overview of the data; the dataset consisted of 128 observations in total, spread over five cities. The number of pumping stations varied largely among cities. In the entire dataset, 53 pumping stations were categorized as 'clean'. Seventy-five pumping stations showed 'severe accumulation of FOG'.

3. Methods

This study focused on quantifying the relationship between catchment demographics (representing FOG disposal patterns), the accumulation of FOG in pump sumps, and whether the pump sump geometry influenced the accumulation of FOG.

A statistical analysis was performed. Instead of a conventional Generalized Linear Model (GLM), a Generalized Linear Mixed Model (GLMM) was applied to account for correlations between the pumping stations that were located in the same city. We applied the procedure as presented in Fig. 2. The procedure consists of four steps: data exploration, model component selection, model selection, and model validation.

3.1. Data exploration

A detailed data exploration was performed. First, relationships between explanatory parameters were investigated. Following the removal of collinear parameters, a GLMM was applied on the remaining dataset. Based on this GLMM, outliers were detected.

3.1.1. Collinearity

Pairwise correlations among explanatory parameters were examined with visual inspection tools and Pearson correlation coefficients (<0.65). In addition, Variance Inflation Factors (VIFs) were used to examine linear dependence among three or more continuous explanatory parameters. A maximum VIF value of 3 was used; more strict than the cut-off range of 5–10 as suggested by Montgomery et al. (1992). One collinear parameter at a time was removed until the values for the VIF and Pearson correlation coefficient were below the preselected thresholds.

3.1.2. Sewer operator dependency

The pumping stations were examined for operator dependency. The data exploration revealed that pumping stations located in the same city showed similarities in their characteristics. As this study aims to identify parameters influencing the accumulation of FOG in pumping stations, revealing the potential effect of unknown city-specific parameters was not in the interest of this research. A GLMM with a random component that accounted for the operator/city effect was therefore applied. This mixed model structure, which is further elaborated in Section 3.2, allows making statements on the relationships for similar cities in general. It describes the notion of an operator and/or city effect, inherently of what comprises such effects.

Table 4
Overview of dataset, showing the total number of observations, and the number of observations with and without FOG accumulation in the pump sump per city.

Municipality	Pumping stations in dataset		
	Total	Clean	FOG
Amsterdam	53	21	32
Rotterdam	12	6	6
The Hague	25	7	18
Almere	26	12	14
Arnhem	12	7	5
Total	128	53	75

3.1.3. Outliers

Based on the GLMM with a random component that accounted for the operator/city effect, and the fixed component containing all parameters that remained after removal of the collinear parameters, the dataset was studied for the presence of outliers. Observations were considered outliers when the severity of FOG accumulation was likely to be caused by industry, and when the simplifications on pumping station geometry and system layout caused a large discrepancy between the actual values and the calculated values.

Since the response parameter is binary and only covers the presence or absence of FOG in the pump sump, there is no possibility for outliers in this parameter.

Outliers in the explanatory parameters were investigated exploiting Cleveland's dot plots, and using Cooks Distance statistics (Cook, 1977). As a Cooks Distance cut off, the value $4(n\text{-}k\text{-}1)^{-1}$ with n for the number of observations and k for the number of regression coefficients was set. The threshold value was used to enhance graphical interpretation, after which the points identified were examined in more detail.

After removal of the outliers, the parameters were checked for collinearity again. The removed outliers did not cause the VIF values and correlation coefficients to rise above the threshold values set.

3.2. GLMM component selection and model selection

Both the GLMM component selection and the model selection (see Fig. 2) were based on the protocol for the top-down strategy for linear mixed models as recommended by Diggle et al. (2002) and applied by Zuur et al. (2009). This protocol suggests starting with a GLMM where the fixed component contains all explanatory parameters. In the second step, the optimal structure of the random component is chosen. This induced a correlation structure between pumping stations that were located in the same city. The third step focuses on obtaining the optimal fixed structure by means of backward selection: the first model contains all explanatory parameters after which the terms are dropped one-by-one, until all terms were significant ($p < 0.05$).

3.2.1. Conditional probability distribution and random component

Conditional on the random effect b_i that accounted for the city-effect of city i where the pumping station j was located, the distribution of the presence/absence of FOG accumulation Y_{ij} is assumed to be binomially distributed with probability $\pi_{ij}|b_i$.

The linear predictor η contains both a fixed and a random component, following the form of the linear regression model:

$$\eta(X_{ij}, Z_{ij}) = \boldsymbol{\beta} \times \boldsymbol{X} + \boldsymbol{b} \times \boldsymbol{Z} \tag{11}$$

where $\boldsymbol{\beta} \times \boldsymbol{X}$ is the fixed component and accounts for the fixed effect, and $\boldsymbol{b} \times \boldsymbol{Z}$ for the random effect. The fixed component is a linear function of the explanatory parameters. $\boldsymbol{\beta}$ is the matrix containing the weights assigned to the explanatory parameters, $\boldsymbol{X}$ is the design matrix of the explanatory parameters.

The random component extends the linear function of the fixed component with a compound symmetrical correlation structure,

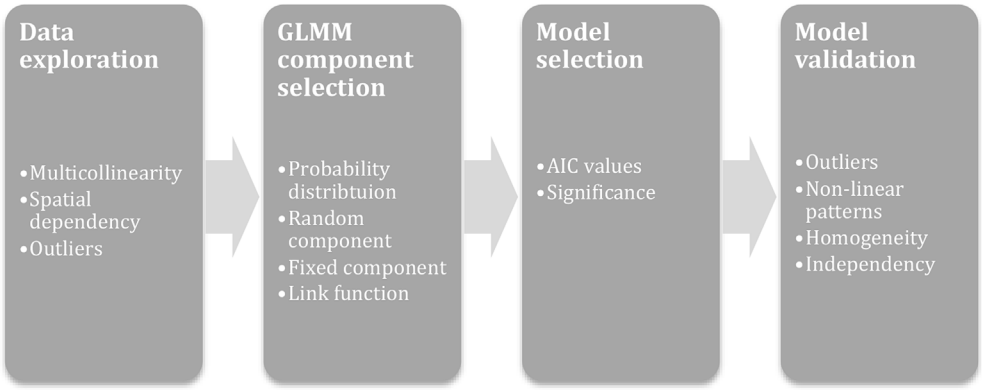

Fig. 2. The procedure for the GLMM model selection and validation process.

adding a random intercept, conditional on city, to the fixed intercept. It models the inter-city variation and assumes that pumping stations that are located in the same city are equally correlated.

3.2.2. Fixed component

The fixed component of the linear predictor is:

$$\eta_{fixed}(X_{ij1}, \ldots, X_{ijM}) = \beta_0 + \beta_1 \times X_{ij1} + \ldots + \beta_M \times X_{ijM} \quad (12)$$

Where j is the pumping station in city i, M represents the total number of explanatory parameters. β_i is the coefficient corresponding to the particular explanatory parameter X, and β_0 is the intercept term.

All continuous explanatory parameters were standardized prior to fitting, facilitating the comparison of the parameter's weights.

3.2.3. Link function

The relationship between the conditional mean and the explanatory parameters is determined by the logistic link:

$$\pi_{ij} = \frac{e^{\eta(\beta_i, X_{ij}, b_i, Z_{ij})}}{1 + e^{\eta(\beta_i, X_{ij}, b_i, Z_{ij})}} \quad (13)$$

3.2.4. Model specification

The final model is:

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \boldsymbol{\beta} \times \boldsymbol{X}_{ij} + b_i \sim N\left(0, \sigma^2\right) \quad (14)$$

where π_{ij} is the probability that FOG accumulates in the pumping station j in city i. $\boldsymbol{\beta}$ is the vector representing the model coefficients, X_{ij} is the vector containing the explanatory parameters for pumping station j, which is located in city i. b_i is the random intercept for city i, and is assumed $N(0,\sigma^2)$.

3.3. Model selection and model validation

A stepwise backwards selection approach was applied to find the optimal model. The assumptions for this final model were verified using visual tools.

The outcome of the GLMM was verified by means of a permutational MANOVA, as implemented in the vegan package (R Core Team, 2016).

4. Results

This section presents the results of the procedure described in Section 3. In the model selection process, nine parameters were dropped (Table 5).

The Cooks Distance statistics designated 15 of the 128 observations as potential outliers. After further exploration of these marked observations, i.e. by inspecting construction drawings and catchment datasheets, two observations were removed as outliers:

- The first pumping station was located in The Hague. More than 25% of the design DWF of this catchment was attributed to industrial wastewater.
- The other station was located in Amsterdam. This pumping station had two inlet pipes, one of which was a pressurized pipe that transported 72% of all incoming wastewater. This specific situation resulted in deviating conditions.

The thirteen remaining marked observations were also checked for particularities in the pump sump geometry and system type. No such particularities were found. Since the high leverage is thought to result from natural variation in pumping stations, and since the Cooks Distance values were still far below the frequently used cut off level of 1, no further observations were removed from the dataset.

Fig. 3 illustrates the differences in a pumping station design philosophy between cities; the conditional boxplots of the kinetic energy density show a larger variation between cities than within cities. For Almere, the median kinetic energy density is $1.6 \cdot 10^6$ [$Jm^{-3}d^{-1}$], which is three orders of magnitude higher than for Amsterdam, where the median value is only $2.5 \cdot 10^3$ [$Jm^{-3}d^{-1}$] (non-log-transformed). In Amsterdam, the construction of most pumping stations is such that they have continuously submerged inlet pipes. This decreases the flow velocity in the inlet pipes considerably and hence, decreases the kinetic energy. In contrast, almost all inlet pipes of pumping stations in the city of Almere are located above the representative water level. This increases the kinetic energy. In addition, the Almere pump sumps are relatively small, which has a positive effect on the kinetic energy per unit of volume and time.

This example illustrates the presence of a city-specific design philosophy, which is supported by the observations made during the data collection and by the authors' knowledge on the Dutch sewer systems. While 'kinetic energy density' is one of the independent parameters in the final model, there could be other (unknown) city-specific parameters influencing the build-up of FOG deposits. It was therefore decided to use a mixed model structure

Table 5
The collinear parameters that were dropped.

Dropped parameter	Reason for dropping
Gutters	Small number of observations (8 in total, 7 of which located in Arnhem
Below social minimum	High number of missing values (21 out of 128)
Personal income (based on working population)	Highest VIF value (48.0) and highly correlated with the parameter 'personal income (based on total population)' (r = 0.96)
Low income population	High VIF value (18.3) and highly correlated with parameter 'renting properties' (r = 0.78)
Non-western immigrants	Highly correlated with the parameters 'housing association properties' (r = 0.67) and 'personal income (based on total population)' (r = -0.69)
Household density	Highly correlated with the parameter 'population density' (r = 0.87)
High income population	Highest VIF value (13.3)
Renting properties	Highest VIF value (11.6)
Daily operation time	Highly correlated with the parameter 'kinetic energy density' (r = 0.78)

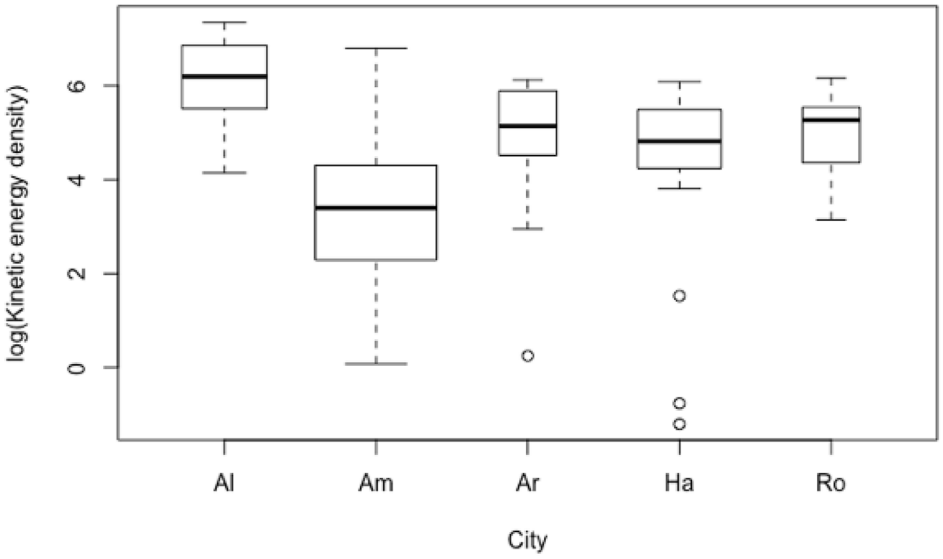

Fig. 3. Boxplots of the total kinetic energy per unit of volume per day (log transformed with base 10), conditional on city. The width of the boxes is proportional to the number of observations per class. The horizontal line in each box is the median, the boxes define the hinge (25–75% quartile). The wide dots represent extreme values.

and deviated slightly from the protocol of Diggle et al. (2002). Incorporating a random effect for city mitigates the potential effect of unknown city-specific parameters, allowing for valid inferences given the available parameters. Such a GLMM structure allows the intercept to be random over cities and assumes a different reference probability for the accumulation of FOG for each city.

Table 6 gives an overview of the model selection process and presents the dropping order of the explanatory parameters. This was based on the relative quality of models as judged by the Akaike's Information Criterion (AIC), and the significance of the model parameters. The parameter that gave the largest drop in AIC if it was excluded from the model, was dropped first. For the final model, the p-values of the estimated regression coefficients should stay stable, i.e. these should not change considerably if one of the parameters is dropped.

During the model selection process, six parameters 'household size', 'population density', 'housing association properties', 'total population', 'sewer system type' and 'vertical flow velocity' were dropped. The final model contains three continuous parameters 'personal income', 'kinetic energy density', and 'FSE density'.

As shown in Table 6, the model with the parameters 'vertical flow velocity' and 'sewer system type' was preferred by the AIC

Table 6
Parameters in the model selection process. The dropping order was based on the significance of regression parameter and the relative quality of the model. The Akaike's Information Criterion (AIC) value is the AIC of the model containing all parameters with a lower position in table. If all model regression parameters were significant, this was indicated with a + in the last column.

Parameter	Type	Dropping order	AIC of GLMM with all parameters below incl.	Significance model parameters (p < 0.05)
Household size	continuous	1	133.62	–
Population density	continuous	2	131.89	–
Housing association properties	continuous	3	130.51	–
Total population	continuous	4	130.07	+
Sewer system type	categorical	5	131.29	–
Vertical flow velocity	continuous	6	132.58	+
FSE density	continuous			
Kinetic energy density	continuous			
Personal income (based on total population)	continuous			

over the final model, and all estimations for the regression parameters were significant. Nevertheless, this model was rejected as the optimal model; after the parameter 'sewer system type' was dropped and the model was fitted again, the regression parameter of 'vertical flow velocity' turned out to be non-significant anymore, making this model not trustworthy.

4.1. Final model

The final GLMM contains the explanatory parameters 'personal income', 'kinetic energy density', and 'FSE density' only. The glmer function from the lme4 package was used for the Bernoulli GLMM, and the model was fit by the default maximum likelihood with a Laplacian approximation. As the GLMM likelihoods involve high order integrals lacking analytical solutions, the likelihood values are approximated using numerical integration.

The final model to estimate the probability of the accumulation of FOG in the pump sump model is specified as:

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = 0.394 - 1.652 \cdot Income_{ij} - 1.068 \cdot Energy_{ij} + 1.749 \cdot FSE_{ij} + b_i \quad b_i \sim N(0, 0.820) \tag{15}$$

where π_{ij} is the probability that FOG accumulates in the pumping station j, which is located in city i.

Table 7 presents the estimated regression coefficients and model fits for this final GLMM with standardized parameters. The probability of FOG accumulation in the pump sump increases in response to a decrease in the average personal income of a catchment area and an increase in the number of restaurants, cafes and cafeterias within a travel distance of 1 km in the catchment area. This probability can be reduced by increasing the daily amount of incoming kinetic energy per unit of volume of water in the pump sump. Table 8 shows the city-specific intercepts. A further discussion on the explanatory parameters and the city-specific intercepts is given in Section 5.2 and 5.3.

The dispersion coefficient, defined as the Pearson residual deviance divided by the residual degrees of freedom in which the mixed effects were calculated to be one degree of freedom, is 0.86. Since this value approximates 1, no over- or under-dispersion could be detected.

The regression parameter estimates are all significant at the 5% level. The parameter 'FSE density' is, however, at the margin of significance with a p-value of 0.0493, using Wald Z-statistics. Comparable results are found when a GLM is fitted as a function of solely the parameter 'FSE density', and gives a (slightly higher) p-value of 0.0661.

As the predictor FSE density was justifiable on the basis of physical considerations, it was decided to keep the parameter in the final model. FOG blockages in sewer lines frequently occur in the proximity of restaurant and bar areas, and most of the FOG deposits analysed were collected from sewer lines downstream from FSE areas (Keener et al., 2008; Shin et al., 2015; Williams et al., 2012). A larger sample size is required to obtain more information on the significance of this relationship.

Table 7
Parameter estimates of the GLMM with standardized parameters. The model estimates the probability of the accumulation of FOG in the pump sump.

Response parameter	Effects		
Y_{ij}	*Random effects*	Variance	
	City identity	0.820	
	Fixed effects	Estimate ± SD	Pr (>|z|)
	Intercept (average)	0.394 ± 0.513	0.4427
	Personal Income	−1.652 ± 0.388	<0.0001
	Kinetic energy density	−1.068 ± 0.468	0.0225
	FSE density	1.749 ± 0.890	0.0493

Table 8
City-specific intercepts and the random effects of the final GLMM.

City	Random effect	Intercept
Arnhem	−0.841	−0.448
Rotterdam	−0.082	−0.475
Amsterdam	−0.007	0.387
The Hague	0.589	0.983
Almere	1.060	1.454

5. Model validation and discussion

5.1. Model validation

Visual tools are used to verify the model assumptions for the final model. Deviance residuals are used for this model validation, enhancing checking for the presence of patterns (McCullagh and Nelder, 1989).

The Cooks Distance statistics is used to check for influential observations once again. No extreme observations were discovered in comparison with the first Cooks Distance plot.

5.1.1. Residual plots

Fig. 4 shows the residuals plotted versus the fitted values, both for all observations at once, and conditional on city. Although residual plots of binomial GLMMs provide only limited information, it is thought that the different cities react comparably to the model.

Fig. 5 shows the deviance residuals against the standardized explanatory parameters for all assessed parameters.

To validate the model, the residual spread should be similar for all values of the explanatory parameter, and no patterns should be present. For the binomial GLMM, the deviance residuals $r_{ij,D}$ are defined as such, that for $Y_{ij} = 0$, $r_{ij,D}$ is negative, and for $Y_{ij} = 1$, $r_{ij,D}$ is positive.

The upper row shows the (standardized) parameters that were included in the final model. For these parameters, the spread was less for higher values, suggesting violation of the homogeneity assumption. Additionally, in the residual plot for kinetic energy density, a pattern can be observed; all residuals are negative for higher values of kinetic energy density.

Most of the parameters that are not included in the model do not show such strong patterns. The parameters 'vertical flow velocity', 'household size', 'housing association properties', and to a certain extent 'sewer system type' displayed residual spreads that are approximately equal for all values of the parameters. Adding the parameters 'population density' and/or 'total population' did not resolve the patterns, nor did adding higher order or interaction terms.

As the patterns could not be resolved, it is concluded that the assumption of independence and constant variance (homogeneity) is violated. This could have affected the estimated regression coefficients.

5.1.2. Permutational MANOVA

To verify the outcomes of the GLMM, a permutational MANOVA, which is more robust to heterogeneity, was applied. A backward selection on the explanatory parameters (Table 6) resulted in a model with the last three parameters equal to the three parameters

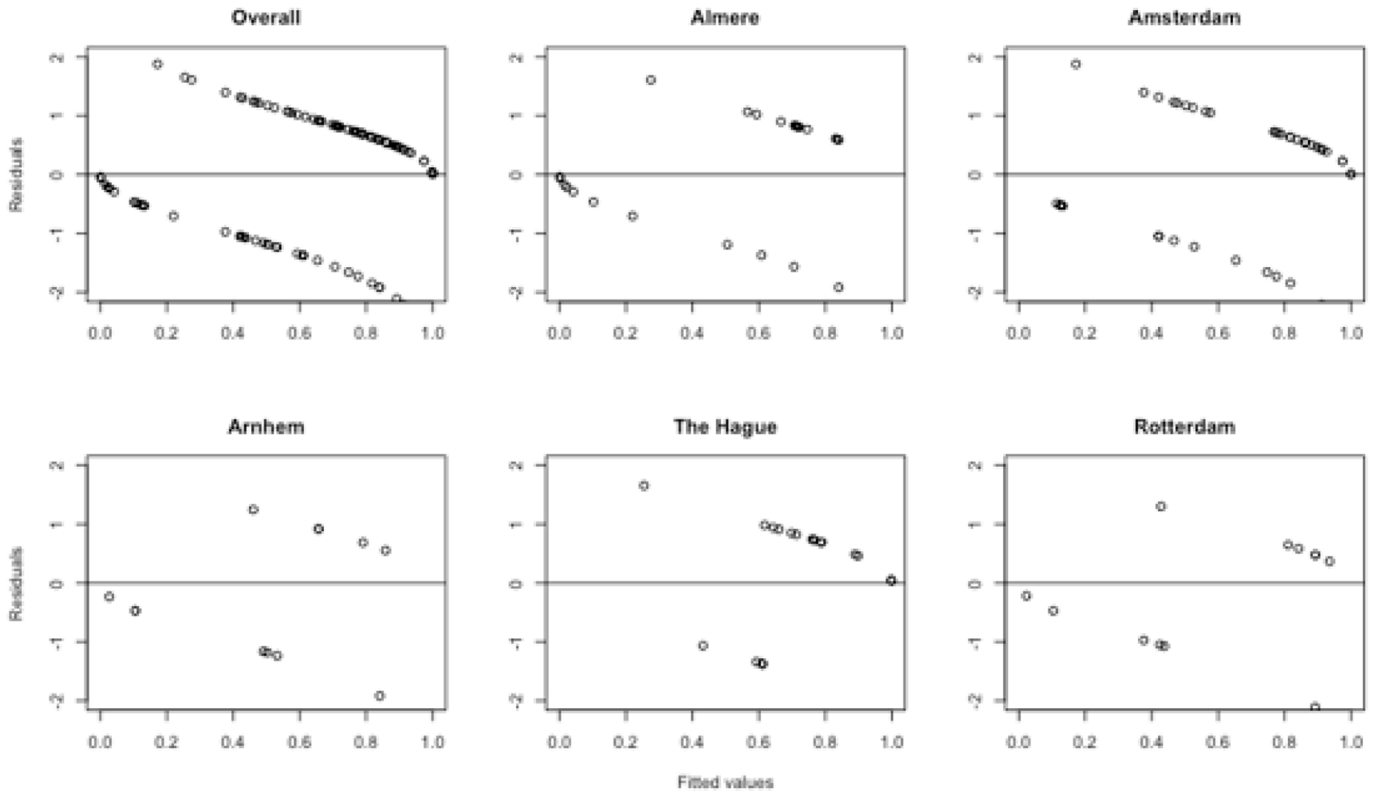

Fig. 4. Deviance residual plots for the fitted values, both for all observations combined and conditional on city. The y-axis shows the residuals, the x-axis the fitted values.

of the final GLMM model ($p = 0.001$), and thereby confirms the outcomes of the GLMM.

5.2. Operator dependency

Fig. 6 illustrates the GLMM predicted probabilities of the accumulation of FOG in pump sumps, along (standardized) personal income values, based on a population mean for the parameters 'FSE density' and 'kinetic energy density'. The thick curve represents the population average, and the two dashed curves represent the inter-city variation; 95% of the values for b_i are estimated to fall between these two curves. The high variance (0.820) reveals that there is a substantial inter-city variation.

To explain the variation between the five participating cities in more detail, a plot of the predicted probabilities of accumulation of FOG per city was made (Fig. 7). This plot shows the predicted probabilities also along the standardized values for personal income, but for the other parameters, the mean values for each city individually were calculated.

The graph illustrates that each city has different intercepts. For a representative pumping station in Arnhem, thus a pumping station with mean values for all parameters for the city of Arnhem, the predicted probability that FOG accumulates in this pumping station is approximately 0.4, while for Amsterdam, this probability equals 0.8.

This shows that pumping stations in Amsterdam are more prone to the accumulation of FOG, given the explanatory parameters. This is also thought to be affected by the relatively low values for kinetic energy density in Amsterdam and high values for the FSE density, making the pumping stations more prone to the accumulation of FOG.

5.3. The role of kinetic energy density and socioeconomic factors related to FOG disposal

The parameter 'kinetic energy density' [$Jm^{-3}d^{-1}$] is the only non-demographic parameter in the model, and its manipulation provides a possible approach to preventing the accumulation of FOG. For example, for catchments with a low average income and a high FSE density, a high kinetic energy per pump sump may prevent the accumulation of FOG in the pump sump.

The significant role that kinetic energy density plays is demonstrated in Fig. 8, showing the probability of FOG accumulation along the standardized parameter for kinetic energy density, for three different income classes. The continuous parameter 'personal income' was discretized into three intervals. The observations were equally divided among the intervals and the mean value for the observations within one interval was taken as the representative interval value.

Fig. 8 demonstrates the importance of kinetic energy for catchments with lower incomes. For a pumping station that is located in a catchment in the low-income class, with a mean value for kinetic energy density (thus the standardized kinetic energy density equals 0), the predicted probability of FOG accumulation is approximately 0.9. For a pumping station located in the same catchment, having different pumping station characteristics, resulting in a value for the standardized kinetic energy density of 4, this probability would be only 0.1. This example illustrates the influence of kinetic energy density on preventing the accumulation of FOG in pump sumps for catchments with a low-income population. For catchments belonging to the high-income class while having an average FSE density, the model suggests that the daily amount of kinetic energy per unit of volume is of less importance. It thereby demonstrates that, besides structural configurations, i.e. the kinetic

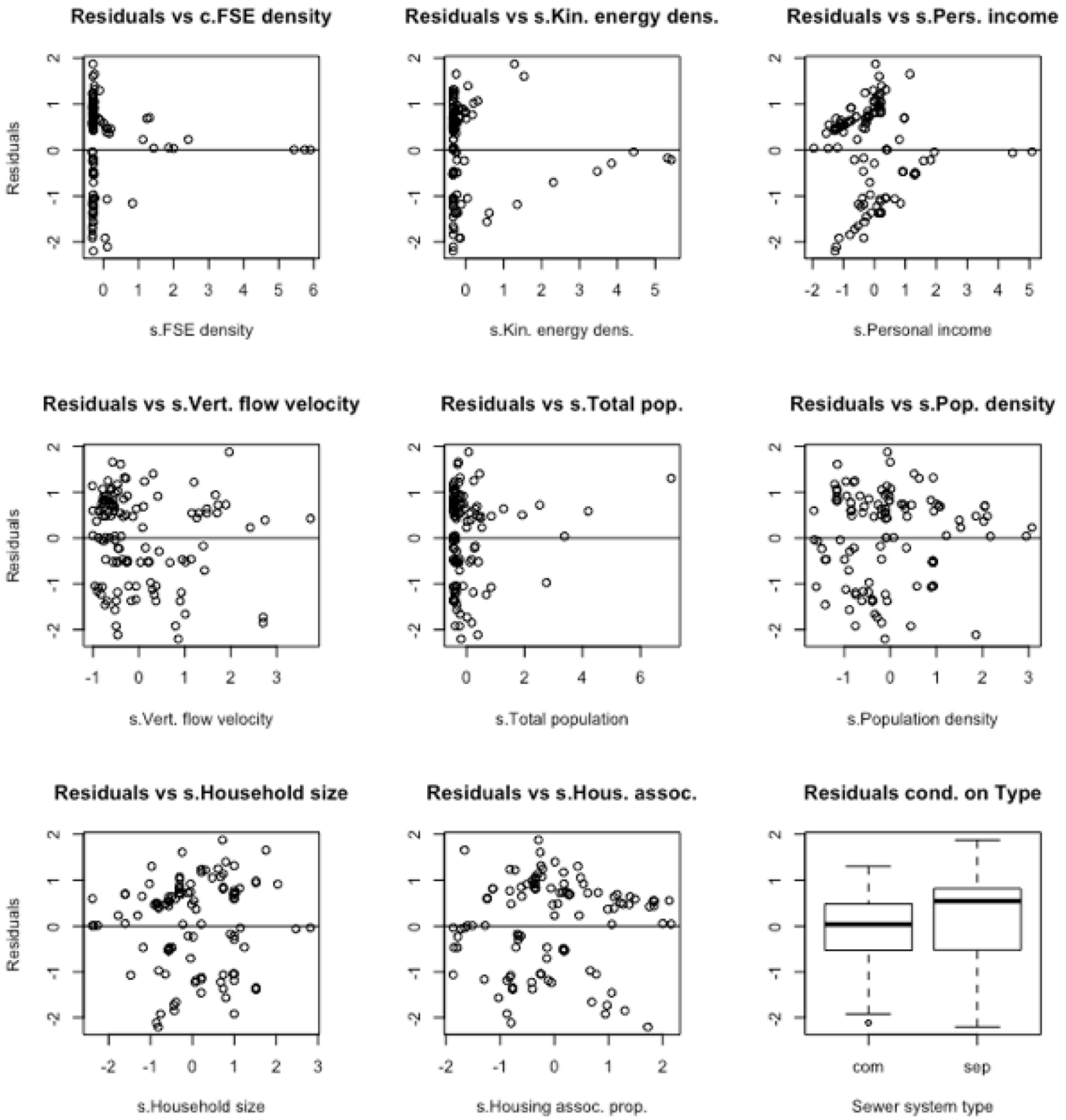

Fig. 5. Residual plots versus the explanatory parameters. The y-axes show the residuals, and the values of x-axis the standardized parameters. The upper row shows the parameters that were included in the final model. The two lower rows show the parameters that were dropped in the model selection process. The line $x = 0$ represents the mean value of the corresponding (standardized) explanatory parameter.

energy density, also demographics, i.e. FOG from FSEs and domestic dwellings, influence the accumulation of FOG.

The high estimated probabilities of FOG accumulation in areas with lower incomes are in line with the observations of Mattsson et al. (2014). They explicitly mentioned the occurrence of severe FOG accumulation in areas with high rise apartment buildings and a relatively high number of immigrants. No such significant relationships could be revealed from this study though; high-rise apartment buildings were not included as such in this study, and population density turned out to be a non-significant parameter. The, on the basis of multi-collinearity, dropped covariate 'percentage of immigrants with non-western origin', however, was highly correlated with the parameter 'personal income' (r = -0.69), suggesting that this covariate may be related to the accumulation of FOG too. Nevertheless, a Dutch governmental study on food habits and lifestyle (RIVM, 2002) seems to contradict this statement: it

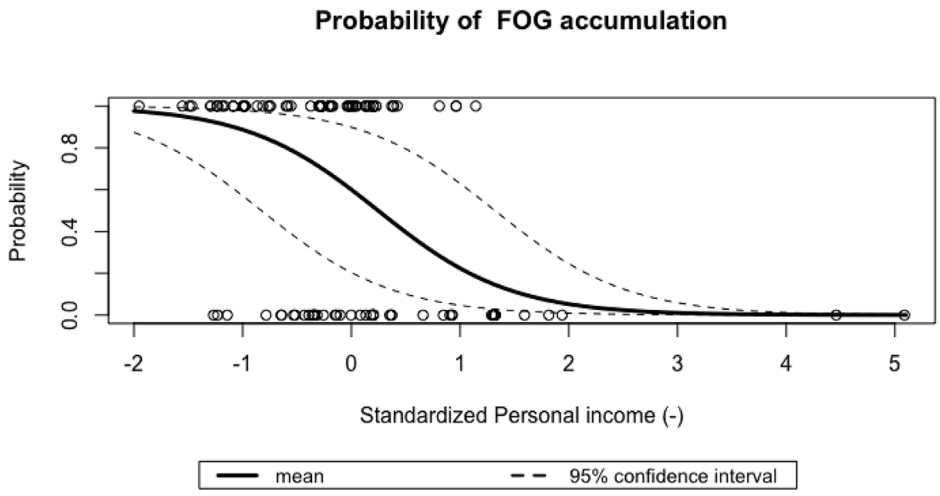

Fig. 6. GLMM predicted probabilities of FOG accumulation along (standardized) personal income values, for catchments with a mean FSE density and for pumping stations with a mean value for kinetic energy density. The thick middle line represents the predicted values for the entire sample of pumping stations. The confidence interval shows the variation of the predictions between the cities.

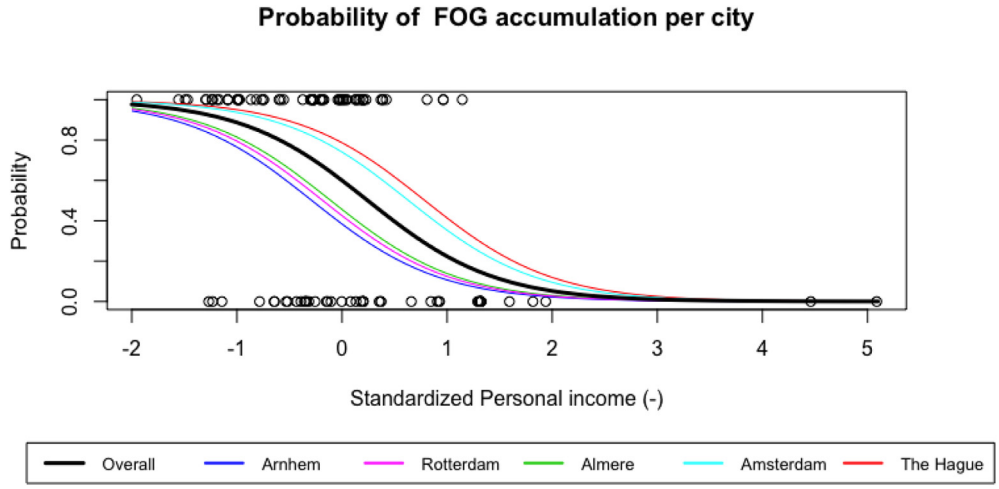

Fig. 7. GLMM predicted probabilities of FOG accumulation per city along (standardized) personal income values, for catchments with both mean values for FSE density and mean values for kinetic energy density of the particular city. The thick middle line represents the predicted values for the entire population.

reported a lower relative fat consumption for Turkish and Moroccan immigrants, both in comparison with groups with low socioeconomic status and with the overall mean of the Dutch population. It should be noted, however, that the Turkish and Moroccan population in the Netherlands represent only around 36% of the entire population that has a non-western origins (Statistics Netherlands, 2016a) and it is unknown whether the Turkish and Moroccan population with low incomes were subtracted from the group with low socioeconomic status. Moreover, FOG disposal is, in addition to fat consumption, also related to cooking and dishwashing habits. As such, the RIVM 2002 study does not allow to conclude on FOG disposal in relation to ethnical groups. Similarly, no such conclusions could be drawn in relation to income. The Dutch National Food Consumption Survey (RIVM, 2011) reported on the intake of fat, subdivided into educational level (and the level of education and the average income are strongly correlated (Statistics Netherlands, 2016b)). This survey revealed only a minor difference (<5%) in the daily fat consumption between the different groups, and moreover, people with a moderate education level had the highest mean fat intake (90.1 g/day). Nevertheless, literature evidence on a relation between income and broader FOG-related issues does exist. A study that analysed differences in the fat intake across social groups for nine European countries, found that people with a lower socioeconomic status consumed slightly more fat than people with a higher socioeconomic status (Lopez-Azpiazu et al., 2003). Another literature review on the geography of fast food outlets found a positive relation between fast food outlets and deprivation (Fraser et al., 2010). Hence, both studies observed a relation between income and issues related to FOG. It is thought that people of one income-group share particular FOG disposal patterns, which could be related to FOG intake and/or cooking and dishwashing habits.

The results of this study suggest that FOG issues in pump sumps may be reduced by minimising the FOG disposal or by increasing the kinetic energy density. Measures to reduce FOG disposal may involve educational campaigns aiming to change the behaviour of people, like the well-known UK 'bin it - don't block it' campaign, or installing grease traps at FSEs AND ascertaining that the grease traps are being operated and maintained properly. Additionally, more robust systems could resolve FOG issues, e.g. if preventive

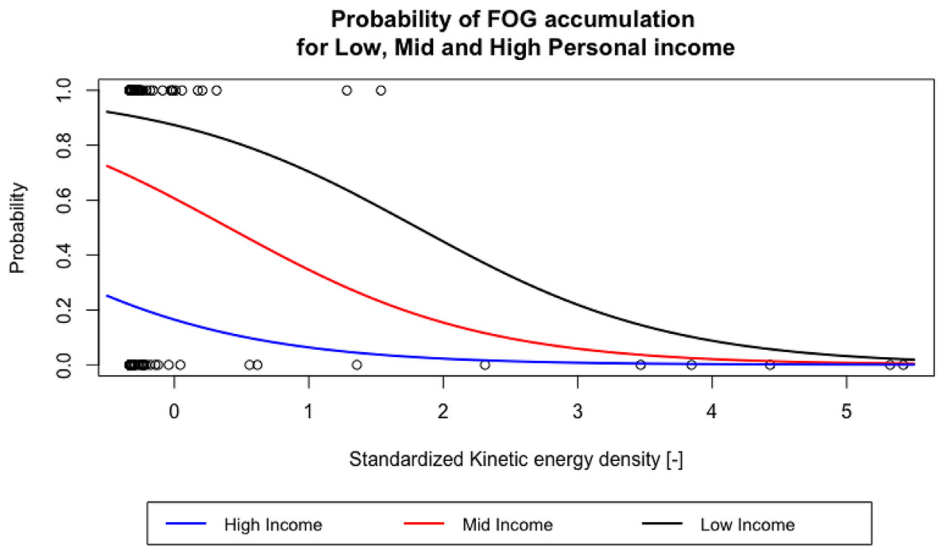

Fig. 8. Predicted probabilities for different income classes for the 'population of cities' ($b_i = 0$), and along (standardized) kinetic energy values. For the parameters 'FSE density' and 'personal income', mean values for the population were taken. The three income classes were based on the intervals of the continuous income parameter. The observations are plotted as dots; the 0 stands for absence and the 1 for presence of FOG. Prediction intervals were not added, as this does not provide valuable information for logistic regression.

measures do not suffice. For pumping stations, FOG accumulation could be overcome by a particular design of pumping stations that accompany higher kinetic energy densities. As this typically requires deeper pump sumps, the design of a pumping station should balance investment costs and operational costs for FOG removal, while at the same time it should avoid excessive air entrainment. From earlier research, see e.g. Lubbers (2007), it is known that the geometry of pump sumps have a large influence on the risk of air-entrainment, which may lead to a significant increase in energy losses in wastewater pressure mains and, in extreme cases, to complete loss of hydraulic capacity (Pothof, 2011; Pothof and Clemens, 2010).

In addition to kinetic energy density, other parameters related to flow velocities and patterns could influence the FOG accumulation process. Dirksen et al. (2012) found that sagging sewers are more vulnerable to the accumulation of FOG, and Dominic et al. (2013) stated that particular sewer constructions, which decrease flow velocities, could enhance FOG accumulation. Furthermore, since the dropped parameter 'daily operation time' was correlated with kinetic energy density ($r = 0.77$), a higher daily operation time could also decrease the probability of the accumulation of FOG. In practice, this suggests, however, operation of pumps beyond the normal operational envelope, which may decrease the service life.

Further physical research on the exact impact of kinetic energy density and flow patterns on FOG accumulation is required. Although the model assumptions of independency and homogeneity were violated, the reported results provide important insights into factors influencing the accumulation of FOG. For future statistical research, it is recommended to systematically record the accumulation of FOG and use a more balanced dataset, i.e. to have more observations with higher values. Also, a larger sample size could solve the observed heterogeneity.

6. Conclusions

This research provides insight into important aspects of catchment demographics and pumping station characteristics that are related to the accumulation of FOG in pumping stations. Generalized Linear Mixed Model (GLMM) procedures are used to analyse the data, consisting of 126 observations of catchments and corresponding pumping stations, located in five different cities. This study presents a procedure to model the probability of the presence or absence of FOG in pump sumps, as a function of demographic and general system characteristics of catchment areas.

The final model contains three parameters, representing the average catchment income, FSE (Food Service Establishments) density, and kinetic energy density of wastewater. The high significance of the parameter 'personal income' demonstrates that it is possible to identify a relationship between FOG disposal and the accumulation of FOG in sewer systems on a catchment scale. This suggests that some aspects of lifestyle, i.e. FOG disposal patterns, are shared by particular demographic groups, thereby resulting in significant variation in the probability of FOG accumulation in pumping stations between catchment areas. Additionally, the analysis shows that geometrical configurations of pumping stations may play an essential role in the prevention of severe FOG accumulation.

The model reveals that severe accumulation of FOG in pump sumps is negatively related to the average income earned per person in the catchments. It is expected that particular FOG disposal patterns are shared by individuals of one income-group, as income cannot influence the accumulation of FOG in itself. Particular diets, cleaning habits and typical moments of FOG disposal might be aspects comprising such disposal patterns, and further research is required to obtain insights into how these aspects may influence the accumulation of FOG. As the dropped parameter 'percentage of non-western immigrants' was highly correlated with income, these particular disposal patterns might be culture-bound.

Furthermore, the model revealed that FSE density is positively correlated with the presence of FOG deposits in pump sumps. As the accumulation of FOG is generally known to be severe in restaurant and bar areas, it is thought that the presence of FSEs directly contributes to the accumulation of FOG.

Next to income and the presence of FSEs, the model finds a negative relationship between the total kinetic energy of DWF per storage volume and presence/absence of FOG in pump sumps.

The results of this study can provide useful information for municipalities in every country to define more effective maintenance strategies or to prevent the accumulation of FOG. It could, e.g., suggest the kind of data that could be recorded by municipalities or motive particular structural configurations of pump sumps. In particular, for catchments receiving wastewater from

areas with a low average income and/or where the FSE density is high, increased construction costs to increase the kinetic energy density may be justified to decrease FOG removal costs. As the assumptions of both independence and homogeneity, however, were violated, the outcomes of the model should be interpreted with care.

For future statistical research, it is recommended to systematically record the accumulation of FOG, use a more balanced dataset and perform (simulation-based) cross-validation to compare model predictions against data. This could improve the predictive performance of the model, thereby providing information for preventing the accumulation of FOG and making municipal maintenance strategies more effective.

The outcomes of this study also provide direction for future experimental design: further research will focus on the multiphase flow phenomena in wastewater pumping stations and on the influence that geometry has with respect to:

- the accumulation of FOG
- air-entrainment
- and sediment deposits

The ultimate goal is to obtain a sound understanding of these processes and to derive a design strategy for wastewater pump sumps that function optimally (e.g. no air entrainment), while their maintenance needs (notably removing FOG and sediments) are minimised.

Acknowledgements

The research is performed within the Dutch 'Kennisprogramma Urban Drainage' (Knowledge Programme Urban Drainage). The involved parties are: ARCADIS, Deltares, Evides, Gemeente Almere, Gemeente Arnhem, Gemeente Breda, Gemeente 's-Gravenhage, Gemeentewerken Rotterdam, Gemeente Utrecht, GMB Rioleringstechniek, KWR Watercycle Research Institute, Royal HaskoningDHV, Stichting RIONED, STOWA, Sweco, Tauw, vandervalk + degroot, Waterboard De Dommel, Waternet and Witteveen + Bos.

The authors would like to especially thank the municipalities that contributed to this research project: Almere, Arnhem, Den Haag, Rotterdam; and the water company of Amsterdam: Waternet.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.watres.2018.02.026.

References

Arthur, S., Crow, H., Pedezert, L., Karikas, N., 2009. The holistic prioritisation of proactive sewer maintenance. Water Sci. Technol. 59, 1385–1396. https://doi.org/10.2166/wst.2009.134.

Ashley, R., Bertrand-Krajewski, J.-L., Hvitved-Jacobsen, T., Verbanck, M., 2004. Solids in Sewers: Characteristics, Effects and Control of Sewer Solids and Associated Pollutants. IWA Publishing, London, UK.

Cammarota, M.C., Freire, D.M.G., 2006. A review on hydrolytic enzymes in the treatment of wastewater with high oil and grease content. Bioresour. Technol. 97, 2195–2210. https://doi.org/10.1016/j.biortech.2006.02.030.

Cook, R.D., 1977. Detection of influential observation in linear regression. Technometrics 19, 15–18.

de Groot, A.C., 2015. A Study on the Fouling Process in Inverted Siphons over Time a Study on the Fouling Process in Inverted Siphons over Time. Delft University of Technology.

Desilva, D., Marlow, D., Beale, D., Marney, D., 2011. Sewer blockage management: australian perspective. J. Pipeline Syst. Eng. Pract. 139–145. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000084.

Diggle, P.J., Heagerty, P., Liang, K., Zeger, S., 2002. Analysis of Longitudinal Data, second ed. Oxford University Press, Oxford, UK.

Dirksen, J., Baars, E., Langeveld, J.G., Clemens, F.H.L.R., 2012. Analysis of fat, oil, and grease deposits in sagging sanitary sewers. In: 9 th International Conference on Urban Drainage Modelling (Belgrade, Serbia).

Dominic, C.C.S., Szakasits, M., Dean, L.O., Ducoste, J.J., 2013. Understanding the spatial formation and accumulation of fats, oils and grease deposits in the sewer collection system. Water Sci. Technol. 68, 1830–1836. https://doi.org/10.2166/wst.2013.428.

EPA, 2004. Report to Congress: Impacts and Controls of CSOs and SSOs. https://doi.org/10.1002/yd.20064. Washington, United States.

Franke, W., Ettl, M., Roldan, D., Kuhn, G., Langholm, A.M., Amar, M., 2011. Fat, oil and grease - sewer contamination prevention strategies and double-dosage concept for fat traps and pressure mains. Water Pract. Technol. 6 https://doi.org/10.2166/wpt.2011.036.

Fraser, L.K., Edwards, K.L., Cade, J., Clarke, G.P., 2010. The geography of fast food outlets: a review. Int. J. Environ. Res. Public Health 7, 2290–2308. https://doi.org/10.3390/ijerph7052290.

Gross, M.A., Jensen, J.L., Gracz, H.S., Dancer, J., Keener, K.M., 2017. Evaluation of physical and chemical properties and their interactions in fat, oil, and grease (FOG) deposits. Water Res. 123, 173–182. https://doi.org/10.1016/j.watres.2017.06.072.

He, X., de los Reyes, F.L., Leming, M.L., Dean, L.O., Lappi, S.E., Ducoste, J.J., 2013. Mechanisms of Fat, Oil and Grease (FOG) deposit formation in sewer lines. Water Res. 47, 4451–4459. https://doi.org/10.1016/j.watres.2013.05.002.

He, X., Iasmin, M., Dean, L.O., Lappi, S.E., Ducoste, J.J., de los Reyes, F.L., 2011. Evidence for fat, oil, and grease (FOG) deposit formation mechanisms in sewer lines. Environ. Sci. Technol. 45, 4385–4391. https://doi.org/10.1021/es2001997.

Keener, K.M., Ducoste, J.J., Holt, L.M., 2008. Properties influencing fat, oil, and grease deposit formation. Water Environ. Res. 80, 2241–2246. https://doi.org/10.2175/193864708X267441.

Lopez-Azpiazu, I., Sanchez-Villegas, A., Johansson, L., Petkeviciene, J., Prättälä, R., Martínez-González, M., 2003. Disparities in food habits in Europe : systematic review of educational and occupational differences in the intake of fat. J. Hum. Nutr. Diet. 16, 349–364.

Lubbers, C., 2007. On Gas Pockets in Wastewater Pressure Mains and Their Effect on Hydraulic Performance. Delft University of Technology.

Mattsson, J., Hedström, A., Viklander, M., Blecken, G.-T., 2014. Fat, Oil, and Grease accumulation in sewer systems: comprehensive survey of experiences of Scandinavian municipalities. J. Environ. Eng. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000813, 140, 04014003.1–04014003.7.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman & Hall, London, UK.

Montgomery, D.C., Peck, E.A., Vining, G.G., 1992. Introduction to Linear Regression Analysis, second ed. John Wiley & Sons.

Nieuwenhuis, E.M., Langeveld, J.G., Clemens, F.H.L.R., 2017. The relationship between Fat, Oil and Grease (FOG) deposits in building drainage systems and FOG disposal patterns. In: 14th IWA/IAHR International Conference on Urban Drainage, pp. 333–340.

Post, J.A.B., Pothof, I.W.M., ten Veldhuis, M., Langeveld, J.G., Clemens, F.H.L.R., 2016. Statistical analysis of lateral house connection failure mechanisms. Urban Water J. 13, 69–80. https://doi.org/10.1080/1573062X.2015.1057175.

Pothof, I., 2011. Co-current Air-water Flow in Downward Sloping Pipes. Delft University of Technology. TU Delft.

Pothof, I., Clemens, F., 2010. On elongated air pockets in downward sloping pipes. J. Hydraul. Res. 48, 499–503. https://doi.org/10.1080/00221686.2010.491651.

QGIS, 2013. Development Team,"Quantum GIS Geographic Information System open source geospatial foundation project.

R Core Team, 2016. R: a Language and Environment for Statistical Computing.

RIVM, 2011. Dutch National Food Consumption Survey 2007-2010 | Part 5 Macronutrients. Bilthoven.

RIVM, 2002. Time for Healthy Living: Health Promotion for Specific Target Groups (In Dutch) (RIVM, Bilthoven. Bilthoven).

Shin, H., Han, S., Hwang, H., 2015. Analysis of the characteristics of Fat, Oil, and Grease (FOG) deposits in sewerage systems in the case of Korea. Desalin. Water Treat. 54, 1318–1326. https://doi.org/10.1080/19443994.2014.910141.

Statistics Netherlands, 2016a. Table on Population: General Figures [WWW Document]. URL. http://statline.cbs.nl/. (Accessed 26 November 2016).

Statistics Netherlands, 2016b. Table on Working Population: Average Income (2005-2012) [WWW Document]. URL. http://statline.cbs.nl/. (Accessed 20 November 2016).

Statistics Netherlands, Kadaster, 2012. District and Neighborhood Map 2012 (In Dutch).

Straub, W.O., 1978. A quick and easy way to calculate critical and conjugate depths in circular open channels. Civ. Eng. 70, 71.

Tukker, M., Kooij, C., Pothof, I.W.M., 2012. Hydraulics Handbook for Pressurised Wastewater Mains (In Dutch), second ed. (Delft, the Netherlands).

Wallace, T., Gibbons, D., O'Dwyer, M., Curran, T.P., 2017. International evolution of fat, oil and grease (FOG) waste management – a review. J. Environ. Manage 187, 424–435. https://doi.org/10.1016/j.jenvman.2016.11.003.

Williams, J.B., Clarkson, C., Mant, C., Drinkwater, A., May, E., 2012. Fat, oil and grease deposits in sewers: characterisation of deposits and formation mechanisms. Water Res. 46, 6319–6328. https://doi.org/10.1016/j.watres.2012.09.002.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R, 2nd Editio. Spring Science and Business Media, New York, New York, US.