Effects of weather conditions on drinking water distribution pipe failures in the Netherlands

B. A. Wols, A. Vogelaar, A. Moerman and B. Raterman

ABSTRACT

The influence of the weather parameters of temperature, wind and drought on pipe failure of drinking water distribution pipes was studied for the Netherlands. Several data sources were used relating weather effects to pipe failure: pipe failure data, regional weather data from different weather stations in the Netherlands, soil settlement data obtained from satellites and (modelled) pressure data. For asbestos-cement (AC) and cast iron (CI) pipes, temperature was an important factor. CI pipes showed increased pipe failures at low temperatures, which confirms results from previous studies, whereas AC pipes showed increased pipe failures at high temperatures. Pipe failure rates were higher for pipes that on average received higher internal pressures. This study also showed that wind resulted in additional pipe failures caused by uprooting of trees during a severe storm. With respect to drought, in some regions in the Netherlands, increased pipe failures during periods of drought were found. A small influence of soil settlement on pipe failure was found using remote-sensing techniques for a small area (5×10 km) in the Netherlands.

Key words | asset deterioration, climate change, drinking water distribution systems, failure database, pipe failure, weather

B. A. Wols (corresponding author)
A. Vogelaar
A. Moerman
B. Raterman
KWR Watercycle Research Institute,
Groningenhaven 7, Nieuwegein 3430 BB,
The Netherlands
E-mail: *bas.wols@kwrwater.nl*

B. A. Wols
Wetsus,
Oostergoweg 9, Leeuwarden 8911 MA,
The Netherlands

INTRODUCTION

Failure of drinking water pipes is caused by various reasons (Le Gat & Eisenbeis 2000; Kanakoudis 2004; Tsitsifli & Kanakoudis 2010). Weather conditions can also influence pipe failure (Newport 1981; Rajani *et al.* 2012; UKWIR 2012; Laucelli *et al.* 2014). As weather conditions are expected to change due to climate change, the number of pipe failures may also change. Temperature and drought are recognized as the most important weather conditions that influence pipe failure (Newport 1981; Rajani *et al.* 1996; Kleiner & Rajani 2002; Rajani & Tesfamariam 2004; Hu & Hubble 2007; Clayton *et al.* 2010; Gould *et al.* 2011; Rajani *et al.* 2012; UKWIR 2012; Wols *et al.* 2013; Wols & Van Thienen 2014a; Pietrucha-Urbanik 2015a, 2015b; Kutyłowska & Orłowska-Szostak 2016). Increased numbers of pipe failures are observed in winter periods (Newport 1981; Rajani *et al.* 1996; Rajani & Tesfamariam 2004; UKWIR 2012; Wols *et al.* 2013; Laucelli *et al.* 2014; Kutyłowska & Orłowska-Szostak 2016), in periods with large temperature differences (Ahn *et al.* 2005; Rajani *et al.* 2012; Pietrucha-Urbanik 2015a) or in summer periods for asbestos-cement (AC) pipes (Wols & Van Thienen 2014a). During periods of drought, higher failure rates are observed, caused by soil shrinkage and/or (differential) soil settlements (Newport 1981; Kleiner & Rajani 2002; Hu & Hubble 2007; Clayton *et al.* 2010; Gould *et al.* 2011). No effect of wind on pipe failure was studied.

In previous work for the Netherlands the influence of temperature on pipe failure was observed (Wols & Van Thienen 2014a): AC pipes showed an increase in pipe failure at high temperatures, whereas cast iron (CI) pipes showed an increase in pipe failure at low temperatures. No effect of wind or drought was found. These results were also translated to predict future pipe failure under different climate-change scenarios (Wols & Van Thienen 2016). This study

doi: 10.2166/ws.2018.085

revealed that the effect of climate change is relatively small, compared with the effect of pipe ageing and pipe replacements. In addition, AC pipes, which are most sensitive to high temperature, are expected to be replaced in the near future. Results were limited to the weather parameter of temperature using data from a single weather station in the Netherlands (De Bilt) for a smaller failure data set (around 40% of the current data set).

Regional differences in weather conditions as well as soil composition and ground water levels occur in the Netherlands. In the current work, a more detailed study for the Netherlands takes place using regional data and more advanced data mining techniques. Also, a longer period of historical failure data was used that included some major storms to assess the effect of wind on pipe failure. Furthermore, explanations of the weather influence on pipe failure are sought by studying the influence of internal pressure and soil settlements using remote-sensing techniques.

MATERIALS AND METHODS

Data sources

The following data sources were used for the analysis: historical pipe failure data and historical weather data.

For pipe failure, data from a Dutch national pipe failure database (USTORE 2009; Vloerbergh & Blokker 2010; Vloerbergh *et al.* 2012) was used. The data consisted of 16,082 pipe and joint failures over a 5-year period (Jan 2009–Dec 2014) and a total pipe network length of 97,667 km (around 80% of all pipes in the Netherlands). The coordinates of all pipes as well as pipe failures were available in a geographical information system (GIS). The following attributes of pipes and/or pipe failures were used: date of failure, cause of failure, pipe material, pipe installation year and pipe diameter. The latter three were also identified by Savic *et al.* (2009) as explanatory variables for pipe failure. The pipes and pipe failures were coupled to a soil type using a soil map of the Netherlands, available in a GIS (WUR 2006). The soil map represents the surficial soil types, representative for pipes typically buried 1 m deep in the ground. For simplicity, four soil types were used: peat, clay, sand and urban land. Failures caused by third parties or failures occurring during installation were removed from the data set. Also, only failures in AC, GCI (grey cast iron), DI (ductile iron), PVC (polyvinyl chloride), PE (polyethylene) and steel (ST) pipes were used. The total number of failures used in the analysis was 12,103.

Regional weather data of different weather stations in the Netherlands was collected from KNMI (Royal Netherlands Meteorological Institute, see Supporting Information). In GIS, Thiessen polygons were created around each weather station. The weather conditions in each of these polygons were assumed to be spatially homogeneous.

The following weather parameters were used in this study: mean daily air temperature (TG, °C), maximum daily wind gust at ground level (FXX, m/s), daily precipitation amount (P, mm) and potential evapotranspiration (E, mm). The effect of wind storm is parameterized by the maximum daily wind gust, as the maximum wind gusts may be responsible for damage to surroundings (e.g. trees). The precipitation and evapotranspiration were used to calculate the rain deficit (RD, mm):

$$RD = \sum_{1}^{n} (E^{(n)} - P^{(n)}) \quad (1)$$

where an index of $N = 1$ refers to 1 April, and N the day of interest. From 1 October until 1 April, the RD was set to zero (defined by the Royal Netherlands Meteorological Institute, KNMI), as in this period in the Netherlands the evapotranspiration is small and no rain deficit occurs.

The pipes (consisting of multiple pipe segments and joints with the same properties) and pipe failures in each weather station polygon were assigned to the particular weather station. If a pipe crossed a polygon, the pipe was cut and divided over the two crossing polygons.

Frequency analysis

The effect of weather conditions on pipe failure was studied by plotting the failure rate against a weather parameter for various cohorts. The weather parameter was therefore divided into ten classes. For each cohort, the failure rate per weather parameter class (PF_i) was calculated, according to:

$$PF_i = \frac{NF_i}{D_i L} \quad (2)$$

In these calculations, only the days (D_i) and number of pipe failures (NF_i) on the days that the weather parameter was within its weather class (i) were considered. L is the total pipe length. Note that a time gap may exist between the date of failure occurrence and discovery (reported in the failure database). Since the trend between weather parameter and failure rate is studied over a number of weather variable classes (Figure 1), it is expected that a failure that due to this time gap falls unfairly into an adjacent weather variable class has a limited influence on the trend between failure rate and weather variable. As the parameters used in Equation (2) may differ per weather station, the failure rate per weather parameter class is calculated for each station. The failure rate for a weather class is then determined by taking the weighted average of all weather stations (weighing over pipe length multiplied by the number of days). An 80% uncertainty bound at the calculated failure rates was determined using a Poisson distribution (based upon the number of failures, see Wols & Van Thienen 2014a).

First, cohorts were made based upon pipe material. In addition, for AC and PVC pipes (that contain most of the failures) cohorts based upon pipe diameter, pipe installation year classes and soil composition were constructed.

Machine learning

Non-linear machine learning techniques were used to relate failure rates with pipe and soil features as well as weather conditions. Only pipe failures were considered, leaving out the joint failures, as the behaviour and cause of joint failures may differ from pipe failures (and less data is available on joint failure, resulting in too small cohorts). Cohorts were constructed based upon pipe material, pipe diameter, pipe year of installation, soil composition and temperature. Only temperature was included as a weather condition (mean daily temperature), as temperature had the largest influence on pipe failure and by adding other weather parameters the number of failures per cohort became too small. The following classes were used for the cohorts: pipe materials AC, PE, PVC, ST, GCI, DI; pipe diameter classes of 0–40 mm, 40–80 mm, 80–150 mm, 150–200 mm, 200–300 mm, 300–600 mm, >600 mm; year-of-installation classes of <1900, 1900–1940, 1940–1960, 1960–1970, 1970–1980, 1980–2000, >2000; soil type classes peat, clay, sand and urban land; and six temperature classes (equally divided between -3 °C and 22 °C). Only the cohorts with more than 4 failures were considered, otherwise the uncertainty of the calculated failure rate of a cohort becomes too large. The choice of these classes is mainly based on occurrence in the distribution network and availability of pipe failure data, so that sufficient cohorts could be obtained to perform the analysis. This resulted in 311 cohorts with a total of 7,123 pipe failures.

For classes with numerical values (diameter, year of installation and temperature), average values of the parameter in each cohort were calculated and used as an explanatory variable in the machine learning. The pipe materials and soil type parameters were converted to categorical variables (variables with a fixed value for each class).

Non-linear regression was used to correlate the explanatory variables to failure rates. This was implemented in the programming language Python using the Scikit-learn toolbox (Pedregosa *et al.* 2011). Different regression techniques were used (gradient boosting regression, random forest, AdaBoost regression, extra trees regression, linear regression), of which the gradient boosting regression (gbr, see Friedman 2001) showed the most promising results and was therefore selected. To reduce the probability of overfitting, the cohorts were divided into a training set of 80% of the cohorts and validation set of 20% of the cohorts. The splitting occurred randomly. Due to the random splitting, results may depend upon the splitting. Therefore, the regression was repeated 100 times to obtain a reliable result of the regression performance.

RESULTS

Influence of weather parameters

The failure rates are plotted against the three weather variables for the different pipe material cohorts (Figure 1).

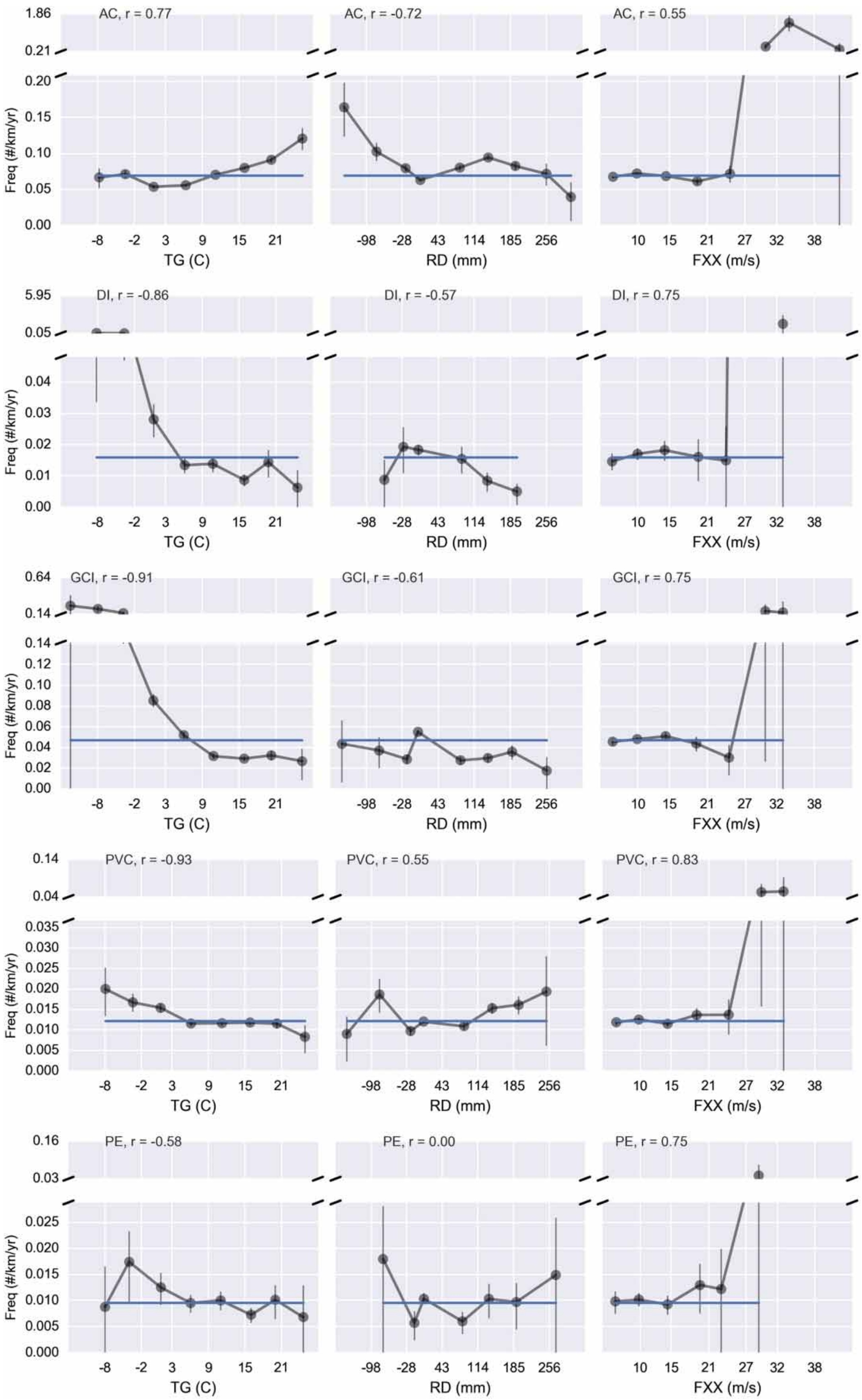

Figure 1 | Relation between weather variable and failure rate (freq) for temperature (left), rain deficit (middle) and wind (right). The error bars show the 80% uncertainty bounds. The *r* shows the Spearman's correlation coefficient between the failure rate and weather parameter.

These cohorts comprise all the diameters, soil types and years of installation. Smaller cohorts that differentiate towards both pipe material and diameter, year of installation or soil type are shown in the Supporting Information.

Temperature

The influence of temperature on pipe failure strongly depends on pipe material (Figure 1). Similar results as presented in previous studies using national weather data (Wols & Van Thienen 2014a) can be observed: failure rates increase with temperature for AC pipes, whereas the opposite occurs for CI, DI and PVC pipes. PE pipes show no significant dependence on temperature. The temperature effect seems to be larger than for AC pipes with smaller diameters (see Supporting Information). This effect was not observed for PVC pipes. The year of installation had no effect on the relation of pipe failure to temperature, both for AC and PVC pipes (see Supporting Information).

Drought

The effect of drought seems to be small: a small increase in failure rate is observed for PVC during periods of high rain deficits, however, the uncertainty is large here. The soil composition also had an effect: for AC pipes and to some extent for PVC pipes in clay and peat soils, the failure rates increase at high rain deficits (see Supporting Information).

Wind

Wind had an evident effect: all pipe materials show a large increase in failure rate at high wind speeds. This was not observed in a previous study in the Netherlands, because no heavy storms occurred in the period of the previous study (2008–2012).

Regional differences

The regional distribution of failure rates is shown in Figure 2, visualized as the increase or decrease in failure rate compared with the average failure rate. Temperature, rain deficit and wind gust are divided into four classes. At high and low temperatures, clear regional differences appear: at low temperatures higher failure rates occur in the northeast and west of the country. The increase in failure rate at high temperature is more pronounced in the east of the country. At the highest rain deficit, some regional differences are also observed, possibly related to the soil composition. For wind, the highest increases in failure rates are observed in the north of the country.

The regional differences were also analyzed for each pipe material (Figure 3). AC pipes in the west showed an increase in failure rates at low temperatures, which was not observed when the data was averaged over the Netherlands. The increase in failure rate for CI pipes at low temperatures occurred over all the country. For PVC pipes, at some locations an increase in failure rate occurred at low temperatures (mainly in the east of the country).

Data mining

Linear correlations (Pearson's r) between the parameters and failure rates show that year of installation and pipe material are correlated with failure rate, but also cross-correlations occur between year of installation and pipe material (Figure 4, left). Smaller years of installation (older pipes) are correlated with higher failure rates. From the non-linear regression, the importance of different parameters on the predicted failure rate can be obtained, showing that pipe material is most important, followed by temperature and year of installation (Figure 4, right). The regression model using failures in all pipes showed a good performance (Table 1). Pipe diameter and soil composition seem to be of little importance compared with temperature or pipe material. For the four most important variables, the regression model calculated the partial dependence on the predicted failure rate (Figure 5). The partial dependence plots show the average partial relationship between a model variable and the predicted variable (failure rate), see also Friedman (2001). In other words, it shows the average trend between a model variable (e.g. temperature, pipe diameter, etc.) and failure rate. For years of installation before the 1950s higher failure rates are observed as well as for the period 1970–1980. Temperature shows a similar relation as observed before in Figure 4: an increase at low temperatures and at high temperatures. Pipe diameter had a smaller influence on pipe failure, a small increase is

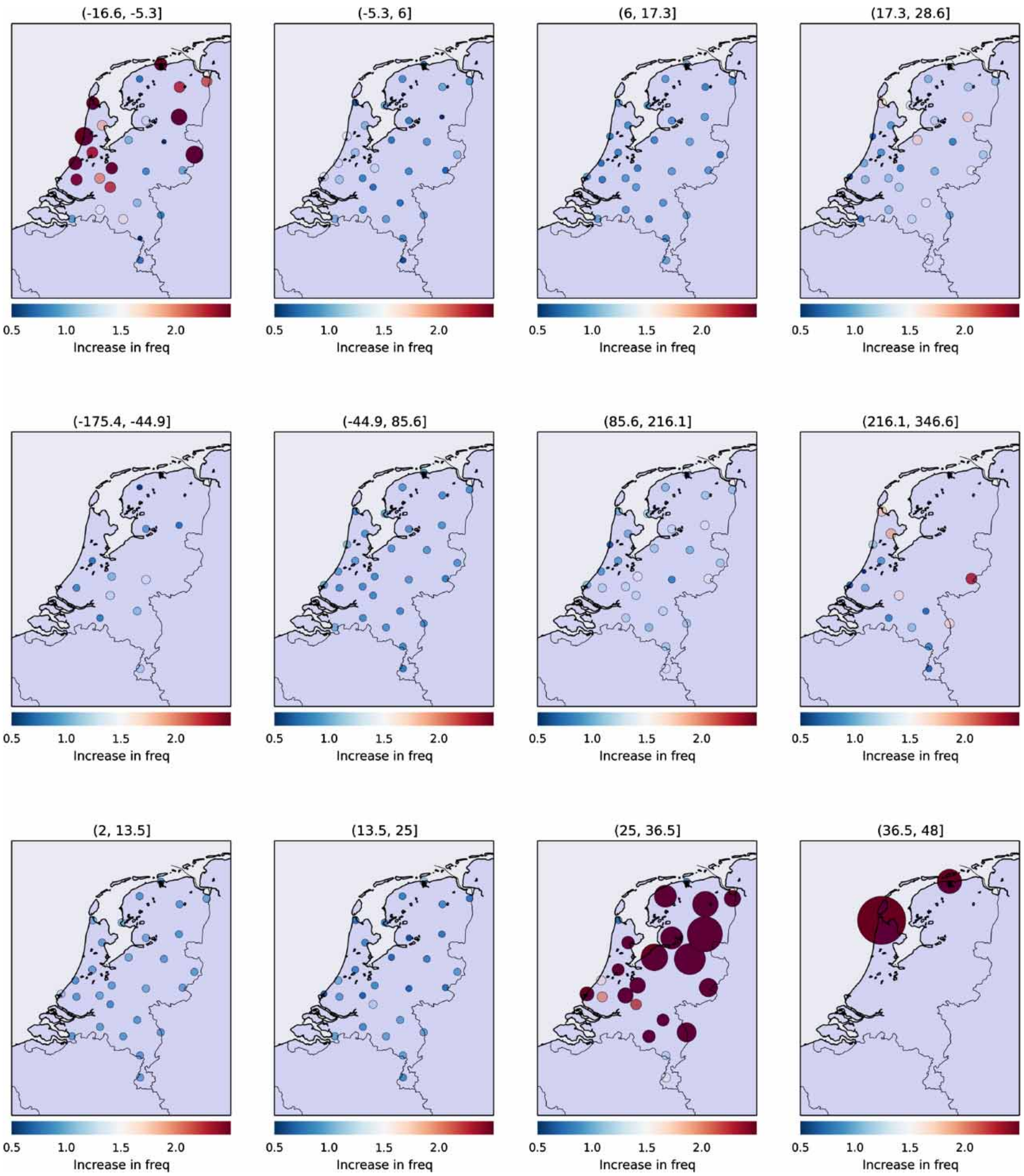

Figure 2 | Regional distribution of the increase in failure rate (freq) for different temperatures (upper), rain deficit (middle) and wind speed (lower) for all pipe materials. Four classes of each weather parameter are shown from left to right.

observed for the larger pipes. This may contradict some findings in the literature, however, most pipes in the Netherlands are PVC pipes, which show higher failure rates at larger pipe diameters (see Supporting Information).

In addition, the regression is conducted for only PVC and only AC pipes. The performance of the prediction for the validation set is relatively low (Table 1), however, some valuable insights into the explanatory variables can be obtained. This is shown in 2D partial dependence plots (see Supporting Information), where the combined effect of temperature and another variable on the predicted failure rate is plotted. For AC pipes, it can be observed that the

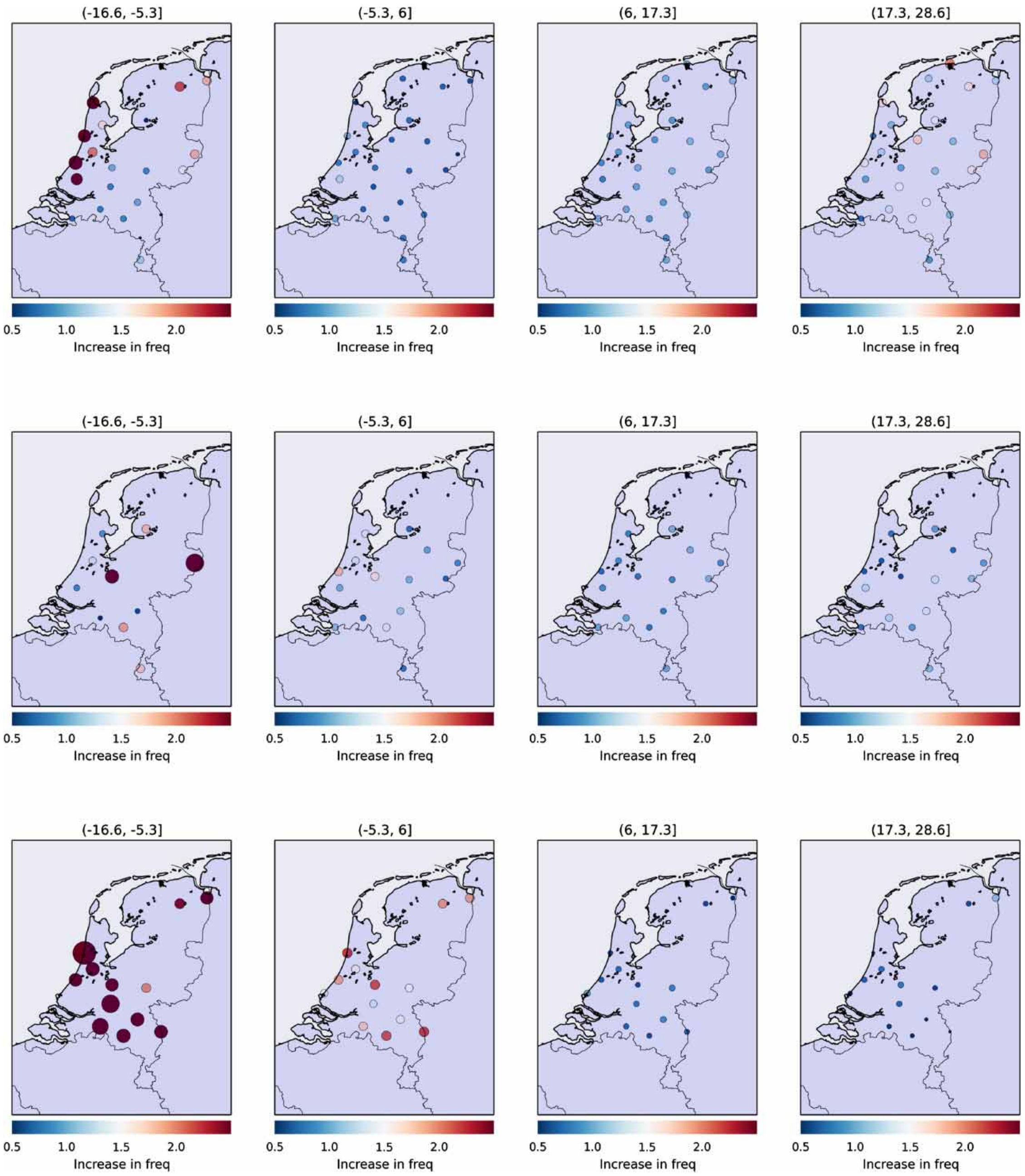

Figure 3 | Regional distribution of the increase in failure rate (freq) for different temperatures in AC pipes (upper), PVC pipes (middle) and GCI pipes (lower).

highest failure rates occur during high temperatures for older pipes, pipes in peat soils or pipes with lower diameters. For PVC pipes, so far little influence of temperature had been observed. From this analysis, some sensitivity of PVC pipes to the lower temperatures can be observed for pipes which are older, or with larger diameters or located in peat soils.

DISCUSSION

The results showed that temperature and severe winds may influence pipe failure, whereas the effect of drought is small. To explain these weather influences, four effects were further analyzed: effect of pressure, effect of freezing, effect of storms and effect of soil settlements.

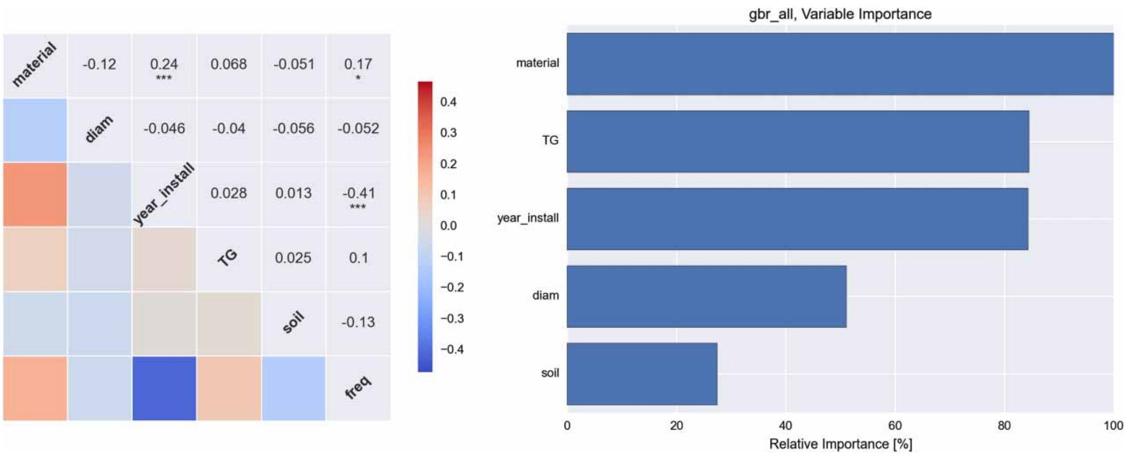

Figure 4 | Correlation between parameters (left) and relative variable importance (right) using the gradient boosting regression statistical model. The variable importance is scaled towards the most important variable, which was set to a value of 100. The stars indicate significant correlations (using Pearson probability value, *: $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Table 1 | Performance of regression model: average goodness of fit after running the statistical model 100 times (R^2)

	Training set	Validation set
All pipes	0.85	0.65
PVC pipes	0.50	0.15
AC pipes	0.75	0.25

Effect of high temperature and relation with pressure (differences)

The effect of temperature in the warmer seasons is less reported in the literature. Ahn *et al.* (2005), Rajani *et al.* (2012) and Pietrucha-Urbanik (2015a) observed pipe failure increase in spring and autumn when large temperature differences occur. In the current study we observe for AC pipe an increase in pipe failure at high temperatures. A possible explanation may be related to pressure. In Wols & Van Thienen (2014a) it was shown that water demand was higher during periods of high temperature. As a result, larger differences in internal pressure over the day and night may occur. A study was therefore performed to relate pipe failures with pressure differences. However, since no spatial distribution of measured pressure over time was available at each pipe (only at some fixed locations in the network), results of pipe network hydraulic models (InfoWorks) were used to obtain a spatial distribution of pressure and pressure differences. The maximum pressure as well as the maximum pressure difference for an average day demand were obtained from the pipe network model. A direct relation between failure, pressure and temperature could not be obtained, as no pressure data at both the location and time was available, but only the pressure for an average day demand at the location of failure. Using these data failure rates for different combined pressure and pressure difference, cohorts could be obtained (Figure 6). The results of individual pressure cohorts and pressure difference cohorts are shown in the Supporting Information. The ranges of the pressure (differences) cohorts were chosen so that the number of failures was equally distributed over the cohorts. In general, failure rates increase when higher internal pressures (for an average day demand) and/or pressure differences (over an average day demand) are applied. Furthermore, the relations between pipe failure and temperature were studied for the different pressure (differences) cohorts. However, the relation of pipe failure with temperature was similar for each pressure cohort (results not shown).

Effect of freezing

Pipe failures also increased at low temperature, especially for DI and GCI pipes, and for some regions in the Netherlands this was also observed for PVC and AC pipes. Other studies also showed increased pipe failures at low temperatures (Newport 1981; Rajani & Tesfamariam 2004). Possible

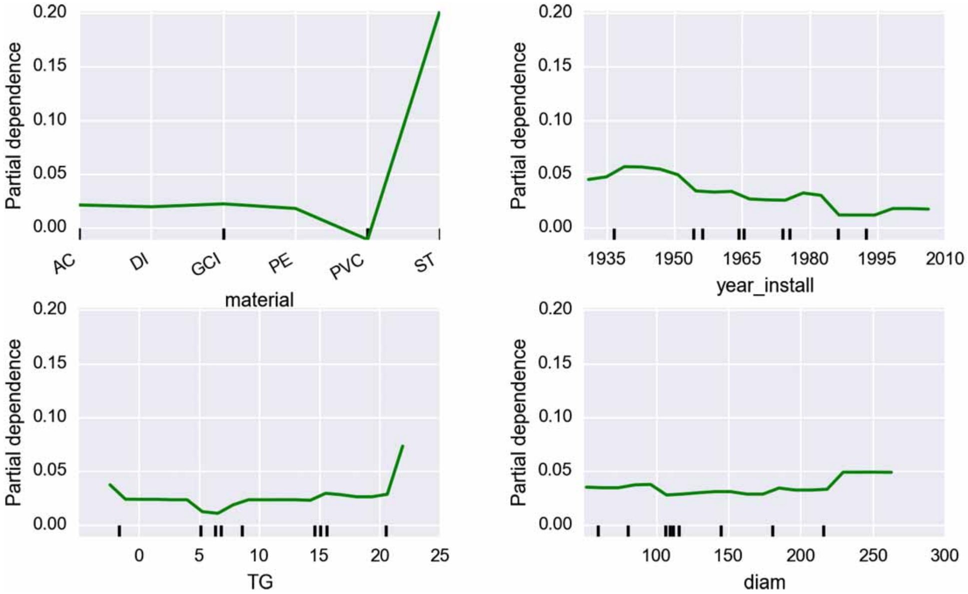

Figure 5 | Partial dependence of different features in the statistical model using all pipes: pipe material (upper left), year of install (upper right), mean temperature (lower left), pipe diameter (lower right). This plot shows the dependence between the predicted target (failure rate) and a set of explanatory variables. The tick marks on the horizontal axis represent the deciles of the variable.

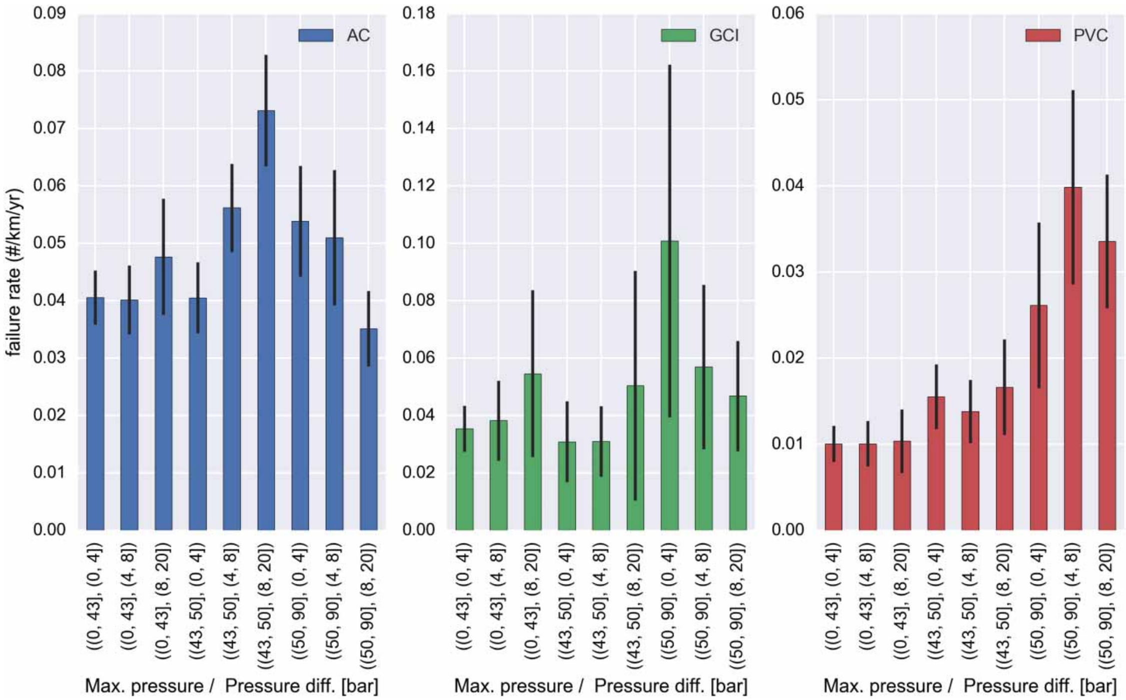

Figure 6 | Effect of maximum pressure and pressure differences for an average day demand on pipe failure (freq) for different pipe materials. Per pipe material, three cohorts of maximum pressure and three cohorts of maximum pressure differences were defined resulting in nine pressure cohorts.

causes are related to soil movements caused by freezing and thawing of soils, occurrence of ice lenses and leaking of joints for GCI pipes. The effect was most pronounced in the west and north-east of the country, which can be related to the soil composition (more peat and clay soils in the west and north of the country). This was also observed in the machine learning results, showing that PVC pipes in peat soils have larger failure rates.

Effect of wind storms

In this study a clear effect of wind was observed related to a severe storm that occurred in the autumn of 2013 (with winds of Beaufort force 11 on 28 October and Beaufort force 10 on 5 December). From the failure registrations, for 90% of the failures around these days trees were registered as cause of failure, whereas in a normal registration for around 20% of the failures trees are indicated as cause of failure. Uprooting of trees is therefore an important cause of failure during severe storms. Figure 7 shows the time series of failures and wind speed in the autumn of 2013. During the storm of 28 October there is a strong increase in the number of failures on the day of the storm and the day after. However, in the storm of 5 December, which was a little less severe, no significant increase in pipe failures was observed. Clearly, a storm has to be sufficiently severe to cause massive uprooting of trees, which is more likely to occur in early autumn, as more leaves are present on the trees. This is also confirmed by the spatial distribution of failures: most of the failures occurred in the north of the country, where the highest wind speeds occurred during the storm. In addition, the most vulnerable and weakest trees were removed in the first storm, so that the number of uprooted trees was probably less in the second storm.

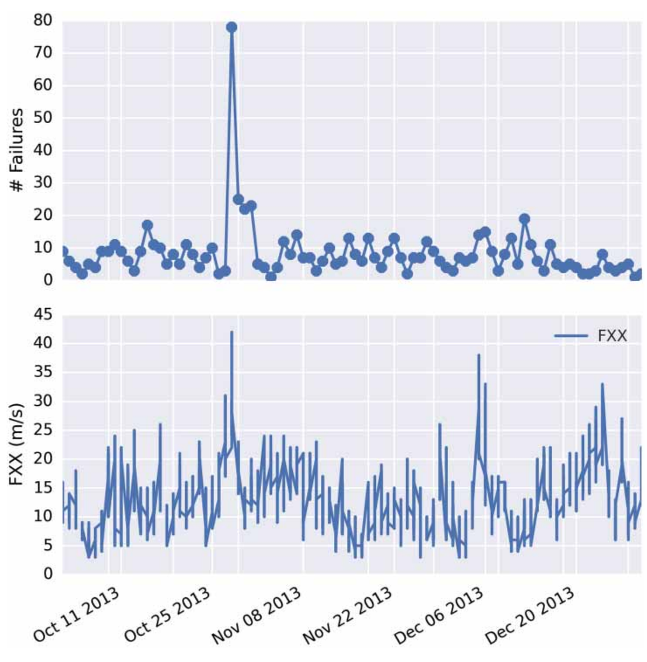

Figure 7 | Time series of pipe failure (upper) and maximum wind speed (lower).

Effect of soil settlements using remote sensing

Soil settlements may occur during warm and dry periods. For the Netherlands, it is expected that soil settlements may increase with climate change up to an additional 10–30 cm by 2050 in the north and west of the Netherlands (Lange & Gunnik 2011). Soil settlements can be measured at a high spatial resolution (3 m × 3 m) using remote-sensing techniques, from which information on soil differential settlements can be obtained. Particularly, PS-InSAR techniques show promising results for ground deformation at a high resolution (Hanssen 2001). Reliable soil deformation data from satellites can be obtained for the so-called persistent scatter points, points of which the scatter characteristics remain consistent in time (across several satellite images). Satellite images from the TerraSAR-X satellite (spatial resolution of 3 m × 3 m) were used to obtain soil deformations for a small region in the west of the Netherlands. Satellite data was collected every 11 days over the period 2009–2014, from which a deformation velocity (in mm/year) was determined by assuming a linear deformation in time. The measured points above the ground (e.g. buildings) were filtered using a digital terrain map (AHN2 2014). Furthermore, the scatter points were interpolated using a nearest-neighbour technique to a 1 m × 1 m map. From this settlement map, a differential settlement map was constructed by taking the maximum slope between the settlement of a cell and its adjacent cells. The settlements were coupled to the individual pipes. For each pipe, the mean soil settlement as well as the mean slope of the soil settlement was determined. The mean soil settlement along the pipe should not be confused with a uniform soil settlement: it shows the mean settlement of the pipe in time, and spatial differences along the pipe can still occur. The mean slope represents the amount of differential settlement that has occurred along the pipe. Soil type is also an important factor, as it determines how large the stresses in the pipe induced by differential soil settlements are (Wols & Van Thienen 2014b), however, the observed region was too small to differentiate between different soil types. The soil settlement parameters were plotted against the failure rate (Figure 8), calculated from 128 pipe failures in the region for which the satellite data was available. There is a slight influence of soil settlement on failure: at higher and lowest (ground uplift) mean deformation failure rates are larger.

CONCLUSIONS

The influence of the weather parameters of temperature, wind and drought on pipe failure was studied for the Netherlands. For AC and CI pipes, temperature was an important factor. CI pipes showed increased pipe failures at low temperatures, which confirms results from previous studies, whereas AC pipes showed increased pipe failures at high temperatures. No direct relation with internal pressure could be found, possibly due to a lack of available pressure data, however, pipe failure rates were higher for pipes that on average received higher internal pressures or pressure differences.

This study also showed that wind resulted in additional pipe failures caused by uprooting of trees. Also, in some regions in the Netherlands increased pipe failures during periods of drought were found. A small influence of soil settlement on pipe failure was found using remote-sensing techniques for a small area in the Netherlands.

In the Netherlands, it is expected that climate change results in higher temperatures (both winter and summer) as well as longer periods of drought (van den Hurk *et al.* 2014), whereas no more severe storms are expected. In a previous study, it was shown for the Netherlands that the expected increase in pipe failure by the higher temperature is small, especially if AC pipes are being replaced by PVC. The current study shows that pipe failure can be further reduced by decreasing the pressure difference during the day and night. The most important threat to the pipe distribution network with respect to climate change may be the expected soil settlements.

In this study, advanced data mining techniques were used as well as satellite data for soil settlements. The results give a first direction on the parameters that contribute to pipe failure, but the uncertainty may still be large and can be reduced when more failure data becomes available. The same also applies to studying the effect of soil settlements. By using a larger area, the

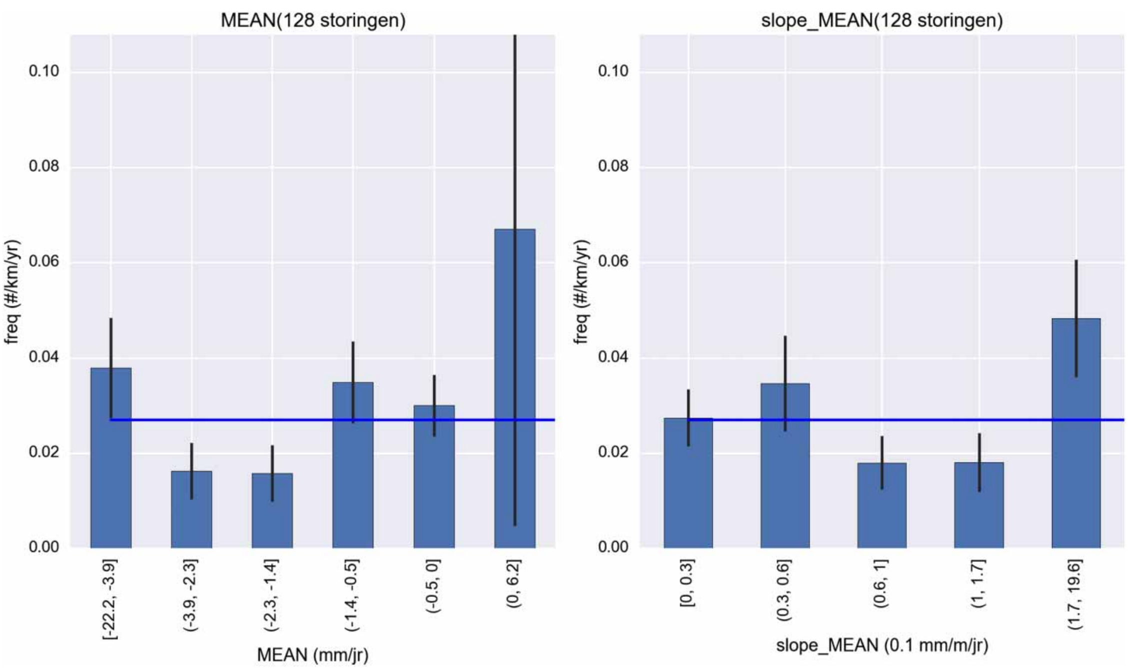

Figure 8 | Failure rate (freq) as a function of soil settlement (left, a negative value is settlement, a positive value is uplift) and mean soil differential settlement (right).

influence of pipe characteristics, such as pipe material, age and diameter with respect to soil settlements could be determined.

The knowledge of the causes of pipe failure can be used to identify weak pipes and prioritize pipe replacements to assist water utilities in support management of their distribution network (Le Gat & Eisenbeis 2000; Kanakoudis & Tolikas 2001; Kanakoudis & Tsitsifli 2011; Pietrucha-Urbanik 2013; Pietrucha-Urbanik 2015b). Results of the statistical analyses could also be coupled with mechanical models that predict mechanical stresses in pipes based upon actual forces (internal pressure, soil, traffic, differential settlements, e.g. Wols & Van Thienen 2014b).

ACKNOWLEDGEMENTS

This study was carried out in the framework of the joint research program of the Dutch Water Utility sector (http://www.kwrwater.nl/BTO). The authors would like to thank the Dutch water companies for the use of failure database USTORE and in particular Brabant Water and WML for providing (modelled) pressure data and Oasen for providing satellite data and pipe failure data.

REFERENCES

Ahn, J. C., Lee, S. W., Lee, G. S. & Koo, J. Y. 2005 Predicting water pipe breaks using neural network. *Water Science and Technology: Water Supply* **5** (3–4), 159–172.

AHN2 2014 *Actueel Hoogtebestand Nederland*, http://www.ahn.nl (accessed September 2017).

Clayton, C. R. I., Xu, M., Whiter, J. T., Ham, A. & Rust, M. 2010 Stresses in cast-iron pipes due to seasonal shrink-swell of clay soils. *Proceedings of the Institution of Civil Engineers: Water Management* **163**, 157–162.

Friedman, J. H. 2001 Greedy function approximation: a gradient boosting machine. *Annals of Statistics* **29**, 1189–1232.

Gould, S. J. F., Boulaire, F. A., Burn, S., Zhao, X. L. & Kodikara, J. K. 2011 Seasonal factors influencing the failure of buried water reticulation pipes. *Water Science and Technology* **63**, 2692–2699.

Hanssen, R. F. 2001 *Radar Interferometry: Data Interpretation and Error Analysis*. Springer Science & Business Media, Dordrecht, The Netherlands.

Hu, Y. & Hubble, D. W. 2007 Factors contributing to the failure of asbestos cement water mains. *Canadian Journal of Civil Engineering* **34**, 608–621.

Kanakoudis, V. K. 2004 A troubleshooting manual for handling operational problems in water pipe networks. *Journal of Water Supply: Research and Technology – AQUA* **53** (2), 109–124.

Kanakoudis, V. K. & Tolikas, D. K. 2001 The role of leaks and breaks in water networks: technical and economical solutions. *Journal of Water Supply: Research and Technology – AQUA* **50** (5), 301–311.

Kanakoudis, V. K. & Tsitsifli, S. 2011 Water pipe network reliability assessment using the DAC method. *Desalination and Water Treatment* **33** (1–3), 97–106.

Kleiner, Y. & Rajani, B. 2002 Forecasting variations and trends in water-main breaks. *Journal of Infrastructure Systems* **8**, 122–131.

Kutyłowska, M. & Orłowska-Szostak, M. 2016 Comparative analysis of water-pipe network deterioration – case study. *Water Practice and Technology* **11**, 148–156.

Lange, G. & Gunnik, J. L. 2011 *Bodemdalingskaarten*. Deltares, Delft, The Netherlands.

Laucelli, D., Rajani, B., Kleiner, Y. & Giustolisi, O. 2014 Study on relationships between climate-related covariates and pipe bursts using evolutionary-based modelling. *Journal of Hydroinformatics* **16**, 743–757.

Le Gat, Y. & Eisenbeis, P. 2000 Using maintenance records to forecast failures in water networks. *Urban Water* **2**, 173–181.

Newport, R. 1981 Factors influencing the occurrence of bursts in iron water mains. *Water Supply and Management* **3**, 274–278.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. 2011 Scikit-learn: machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830.

Pietrucha-Urbanik, K. 2013 Multidimensional comparative analysis of water infrastructures differentiation. In: *Environmental Engineering IV* (A. Pawlowski, M. R. Dudzinska & L. Pawlowski, eds), CRC Press, Boca Raton, FL, USA, pp. 29–34.

Pietrucha-Urbanik, K. 2015a Failure analysis and assessment on the exemplary water supply network. *Engineering Failure Analysis* **57**, 137–142.

Pietrucha-Urbanik, K. 2015b Failure prediction in water supply system – current issues. In: *Theory and Engineering of Complex Systems and Dependability*, Vol. 365 (W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak & J. Kacprzyk, eds). Advances in Intelligent Systems and Computing, Springer-Verlag, Heidelberg, Germany, pp. 351–358.

Rajani, B. & Tesfamariam, S. 2004 Uncoupled axial, flexural, and circumferential pipe–soil interaction analyses of partially supported jointed water mains. *Canadian Geotechnical Journal* **41**, 997–1010.

Rajani, B., Zhan, C. & Kuraoka, S. 1996 Pipe–soil interaction analysis of jointed water mains. *Canadian Geotechnical Journal* **33**, 393–404.

Rajani, B., Kleiner, Y. & Sink, J.-E. 2012 Exploration of the relationship between water main breaks and temperature covariates. *Urban Water Journal* **9**, 67–84.

Savic, D. A., Giustolisi, O. & Laucelli, D. 2009 Asset deterioration analysis using multi-utility data and multi-objective data mining. *Journal of Hydroinformatics* **11** (3–4), 211–224.

Tsitsifli, S. & Kanakoudis, V. 2010 Predicting the behavior of a network pipe using the 'critical *Z*-score' as its performance indicator. *Desalination* **250** (1), 258–265.

UKWIR 2012 *Impact of Climate Change on Asset Management Planning*. Report No. 12/CL/01/16, UKWIR.

USTORE 2009 *Dutch National Failure Database*, www.ustoreweb.nl (accessed 1 May 2018).

van den Hurk, B., Siegmund, P. & Klein Tank, A. 2014 *KNMI'14: Climate Change Scenarios for the 21st Century – A Netherlands Perspective*. KNMI, De Bilt, The Netherlands.

Vloerbergh, I. & Blokker, M. 2010 Sharing failure data to gain insight into network deterioration. *Water Asset Management International* **6**, 9–14.

Vloerbergh, I., Schipdam, R., Thienen, P. & Beuken, R. 2012 Pipe fitters assist to predict investment needs for water main rehabilitation. *Water Asset Management International* **8**, 12–18.

Wols, B. A. & Van Thienen, P. 2014a Impact of weather conditions on pipe failure: a statistical analysis. *Journal of Water Supply: Research and Technology – AQUA* **63**, 212–223.

Wols, B. A. & Van Thienen, P. 2014b Modelling the effect of climate change induced soil settling on drinking water distribution pipes. *Computers and Geotechnics* **55**, 240–247.

Wols, B. A. & Van Thienen, P. 2016 Impact of climate on pipe failure: predictions of failures for drinking water distribution systems. *European Journal of Transport and Infrastructure Research* **16**, 240–253.

Wols, B. A., Van Daal, K. & Thienen, P. 2013 Effects of climate change on drinking water distribution network integrity: predicting pipe failure resulting from differential soil settlement. In: *12th International Conference on Computing and Control for the Water Industry, CCWI2013*, Perugia, Italy.

WUR 2006 Grondsoortenkaart, Wageningen University, http://www.wur.nl/en/show/Grondsoortenkaart.htm (accessed September 2017).

First received 6 October 2017; accepted in revised form 16 April 2018. Available online 30 April 2018