Temperature-dependence in sewer blockage frequency

Josh Nightingale, Christian Gunninng and Mark Holstad

October 15, 2014

1 Methods

1.a Data

Albuquerque Bernalillo County Water Utility Authority (ABCWUA) responds to SSB events after discovery by maintenance workers or reports of blockages from the public. This study used an anonymised dataset of SSB dates, along with engineers' estimates of blockage cause. SSB reports from the period January 2009 - January 2013 (inclusive) were used in this study. The frequency of sanitary sewer blockages is the primary focus of this work.

Sewer grab sample temperature (SGST) measurements were collected by ABCWUA personnel during routine maintenance, using XXX probes XXXX etc. (??mark) SGST measurements were available from 15 manholes, leading to 3 interceptors within Albuquerque. ??(How many total samples) ??(Mark - explain interceptors / structure of this data) SGST from the period December 2005 - December 2010 (inclusive) were used in this study.

Mean daily air temperature (MDAT) was obtained from the Albuquerque International airport's (KABQ) automated METAR data collection system (available from `http://www.wunderground.com/history/airport/KABQ`) for the period of record for which either SSB and SGST or measurements were available (December 2005 - January 2013, inclusive).

Since most days had no SSB events, the total number of SSB events per week (W-SSB) was computed and used in subsequent analysis. For comparison with SSB data, MDAT measurements were averaged by week to yield mean weekly MDAT (MW-MDAT). In addition, SGST measurements were averaged by week (all interceptors were combined), yielding mean weekly SGST (MW-SGST). Weeks without SGST measurements were excluded from any subsequent analyses that included SGST.

1.b Linear models

All analysis was conducted with the R statistical programming environment (R Core Team 2013). For final selection of linear model specifications, both Bayes'information criterion (BIC) and parsimony considerations were employed.

First, we used ordinary linear models to estimate the response of MW-SGST to MW-MDAT, interceptor identity, and manhole identity. In favor of parsimony, and due to the small effect sizes and/or statistical non-significance, interceptor and manhole identity were excluded from subsequent models. Next, we used generalized linear models (GLM) to estimate the response of W-SSB to either MW-SGST or MW-MDAT. We also used a GLM to estimate the response of W-SSB to both MW-MDAT and blockage cause. Due to low sample numbers, the response of W-SSB to both MW-SGST and blockage cause was not estimated.

2 Results

2.a Air temperature (MW-MDAT) and sewage temperature (MW-SGST)

We found that sewage temperature increased with air temperature (Figure 1). Indeed, the best-ranked model of the response of MW-SGST to MW-MDAT (Table 6) explained the majority of variation in MW-SGST ($R^2 = 0.778$). However, as air temperature fells below freezing, little further decrease in sewage temperatures was observed (Figure 1).

For reference, all candidate models (ranked by BIC) are shown in Table 5. All high-ranked models (low BIC) show a small but statistically significant effect of interceptors identity on sewage temperature. On the other hand, all models that included manhole identity ranked lower than the null model (which included only MW-MDAT and MW-SGST).

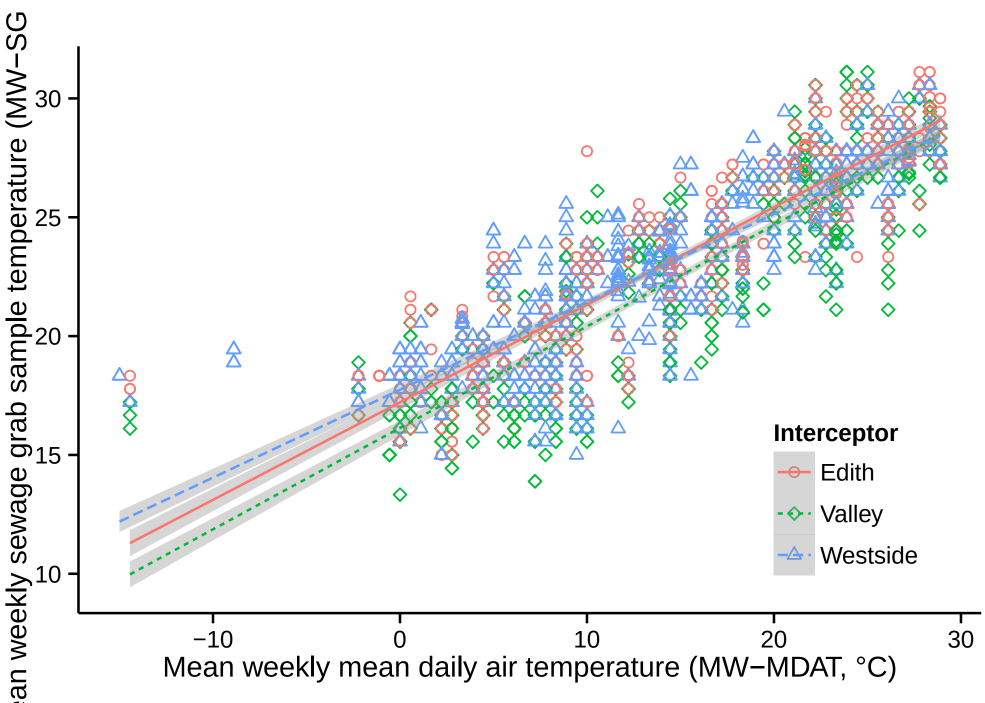

Figure 1: Sewage temperature increased with air temperature, though interceptors differed slightly in their sewage's temperature responses to air temperature. As air temperature dropped below freezing, no further decrease in sewage temperature was observed.

2.b Sewage temperature (MW-SGST) and sewer blockage frequency (W-SSB)

We find that sewer blockages occured more frequently during weeks with lower sewage temperatures (Figure 2A). The final linear model specification of the response of W-SSB to MW-MDAT employed a negative binomial GLM with a log link function Venables and Ripley 2002. The final relationship was highly statistically significant ($p < 0.001$; Table 1),

The W-SSB data followed an overdispersed Poisson distribution, with forty-one weeks (23.8%) showing one or zero incidents. Consequently, the final negative binomial specification provided a significant improvement over a Poisson GLM (likelihood ratio tests, $p < 0.001$).

R^2 statistics are not available for GLM, though the proportional reduction in deviance (D) provides an analogous measure of the model's explanatory power (**zheng2000summarizing**). For the final negative binomial model, we find that D = 0.059, showing that this model explains only a modest amount of variation in observed sewer blockage frequency.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.9713	0.6547	4.54	0.0000
SewTempC	-0.0900	0.0283	-3.18	0.0015

Table 1: Weekly mean sewage temperature predicts the number of blocked sewers that week ($n = 153$ weeks)

2.c Air temperature (MW-MDAT) and sewer blockage frequency (W-SSB)

We found that sewer blockages were more frequent when air temperature was low. MW-MDAT was a highly significant predictor of W-SSB (Table 2; Figure 2B). However, this model explains very little variation in blockage frequency (D = 0.018). ??is this correct? Text said .04, I'm getting .017.

This dataset includes 491 (??code) blockages where grease was the estimated blockage cause, representing 54% of incidents during the study period. When these grease-caused SSB events were modeled separately, a statistically significant relationship between W-SSB and MW-MDAT was also observed (Table 3). ??Fig3 – see greaseplot and nongreaseplot ??Model to test by cause However, there was no relationship between temperature and non-grease SSB (Table 4; Figure 3). For the model predicting grease-caused blockages, $D = 0.040$, compared with $D = 0.005$ for the model predicting non-grease-caused blockages.

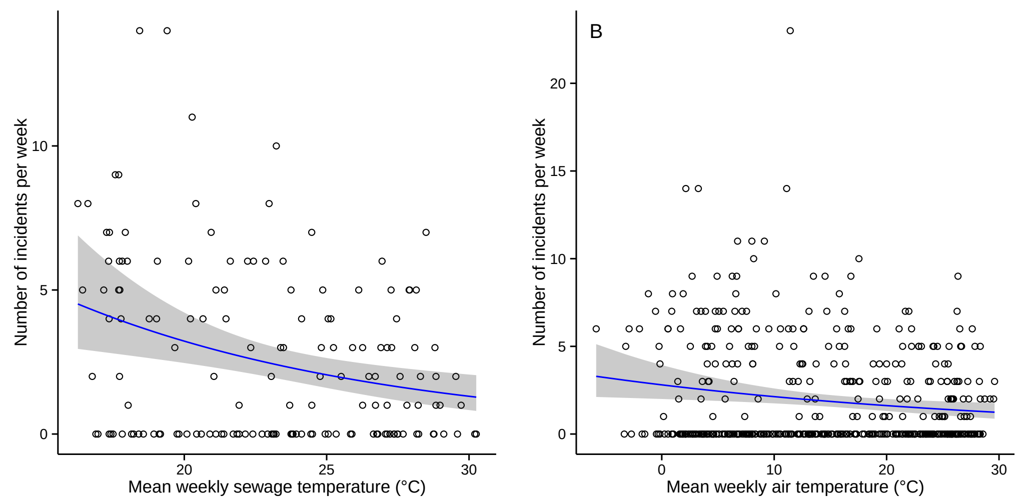

Figure 2: Both sewage temperature (A; $n = 153$ weeks) and air temperature (B; $n = 471$ weeks) can be used to predict sanitary sewer blockages.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.0313	0.1728	5.97	0.0000
MeanTempC	-0.0274	0.0102	-2.68	0.0074

Table 2: Weekly mean air temperature predicts the number of blocked sewers that week ($n = 471$ weeks)

3 Discussion

Notes / todo: ??both sewer temp and air temp are signif predictors, sewer temp is better ??likely that the reponse of sewer temp to air temp is dependent on local local climate and sewer configuration (mark?), warrants testing in different locales. ABQ diurnal temp, elevation gradient and cold

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.5586	0.1616	3.46	0.0005
MeanTempC	-0.0393	0.0099	-3.99	0.0001

Table 3: Negative binomial GLM predicting blockages caused by grease ($n =$ 471 weeks)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.0812	0.1673	0.49	0.6276
MeanTempC	-0.0140	0.0099	-1.41	0.1573

Table 4: Negative binomial GLM predicting blockages not caused by grease ($n = 471$ weeks)

air drainage. ??if municipalities are already collecting SGST, it would be an appropriate addition to system maintainance planning

Temperature data, which are widely and freely available, have modest utility in predicting sewer blockages over weekly timescales. These results suggest that areas experiencing increasing average temperatures may find that this trend alleviates the pressure placed on sewage systems by FOG deposits. Similarly, weather forecasts and real-time weather observations may prove useful for predicting and responding rapidly to blockages, reducing the threat to property and public health.

Data from sewer measurements are a slightly more accurate predictor of blocking frequency. Where these data are regularly collected and rapidly analysed, they could be used in place of air temperature to anticipate problems in sanitation infrastructure.

The relationship between air and sewage temperature is likely to be mediated by ground temperature, and therefore by groundwater levels. The difference in predictive ability between sewage and air temperature may reflect the variable groundwater levels during the seasonal cycle in Albuquerque. Similarly, differences between interceptors and manholes may reflect elevation and land use, via their effects on groundwater temperature. Models including precipitation patterns and/or local physical characteristics (e.g. water table height, land use, soil type, geology) could test this hypothesis.

With continuing population rise and urbanisation, efficient operation of urban waste-water infrastructure is an increasingly important issue for global public health. Sato et al. 2013 recently highlighted the importance of more

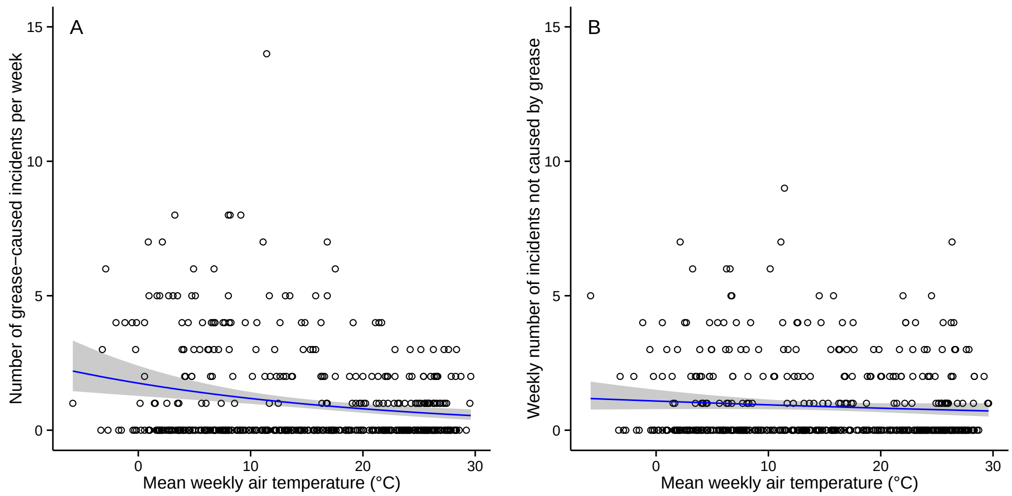

Figure 3: Mean weekly air temperature predicts blockages caused by grease (A; $n = 471$ weeks) but not other blockages (B; $n = 471$ weeks)

research into efficacy of waste-water treatment techniques, particularly in the developing world. The data in this study were not collected specially for research purposes. Rather, this study used data already collected by industry as part of standard operations, married with publicly accessible weather data available online. This demonstrates the potential usefulness of historic industry datasets for addressing future challenges.

References

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. URL: `http://www.r-project.org/`.

Sato, Toshio et al. (2013). "Global, regional, and country level need for data on wastewater generation, treatment, and use". In: *Agricultural Water Management* 130, pp. 1–13. DOI: `10.1016/j.agwat.2013.08.007`.

Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics with S*. Fourth ed. ISBN 0-387-95457-0. New York: Springer. URL: `http://www.stats.ox.ac.uk/pub/MASS4`.

4 Supplemental Information

	Model	BIC
6	mb.by.interceptor	8094.78
2	b.by.interceptor	8105.16
4	m.by.interceptor	8147.29
1	null	8160.33
3	b.by.manhole	8176.95
5	m.by.manhole	8229.28
7	mb.by.manhole	8242.55

Table 5: Candidate models for predicting sewage temperature using mean air temperature, ranked using Bayes Information Criterion

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	17.2	0.163	105.507	0
MeanTempC	0.41	0.0094	43.655	2.69e-289
InterceptorValley	-1.08	0.217	-4.964	7.51e-07
InterceptorWestside	0.529	0.211	2.507	0.0122
MeanTempC:InterceptorValley	0.0165	0.0125	1.322	0.186
MeanTempC:InterceptorWestside	-0.0406	0.0124	-3.267	0.0011

Table 6: Summary table of the best-ranked model of the response of mean weekly sewage grab sample temperature (MW-SGST) to mean weekly mean daily air temperature (MW-MDAT). Sewer interceptor identity has a significant effect on both model slope and model intercept. $R^2 = 0.778$.

```
## Warning:  no non-missing arguments to min; returning Inf
```