Online Bayesian Inference of Diffusion Networks

Shohreh Shaghaghian[1] and **Mark Coates**[1]

[1]Department of Electrical and Computer Engineering, McGill University, Montreal
shohreh.shaghaghian@mail.mcgill.ca , mark.coates@mcgill.ca

In this report, we provide support for the Gaussian model used for the number of weekly reported cases of measles and chickenpox in England and Wales. The analysis is based on the data (shown in Figure 1) from seven large cities (London, Bristol, Liverpool, Manchester, Newcastle, Birmingham, and Sheffield), with populations ranging from $300,000$ to 10 million (out of a total population of approximately 50 million). The primary infections occur every two years in the period from September to December of the following year. We focus our analysis on these time windows. Our model for this data assumes that an infection commences in one city and then by movement of individuals between two cities emerges in one of the other cities. The model does allow for the possibility of infections arising spontaneously due to unrelated influences in multiple cities. In such a case, multiple nodes (cities) have no parent in the inferred model. This can capture scenarios in which infections are caused by interactions with other cities not included in the model (e.g., visitors from other countries). It can also explain the case when the residual infection in a city gives rise to a new outbreak, with no disease transfer from another city.

Quantitative studies (e.g. [1]) have explained the decrease in the number of reported cases in the data after widespread use of a vaccine began in 1968. Also, studies such as [2] have justified the biennial data peaks, claiming that they were caused by exhaustion and subsequent build-up of susceptibles in the population as well as seasonal changes in virus transmission. Finally, [3] shows that once an infection starts in a region, the number of hospitalized cases can be approximated by a log-normal function of the number of days that have passed since the infection began. The number of hospitalized cases in region i at the n^{th} week is denoted by h_n^i and defined as the cumulative number of cases minus the cumulative number of deaths and recoveries. Hence, for $n \geq t_1^i$ we have

$$h_n^i = \frac{A_h^i}{\sqrt{2\pi}\sigma_h^i 7(n-t_1^i)} \exp\left(-\frac{(\ln[7(n-t_1^i)]-\mu_h^i)^2}{2{\sigma_h^i}^2}\right). \tag{1}$$

Here A_h^i is a constant, σ_h is the variance parameter of the log-normal, and we set $\mu_h^i = \ln D_h^i + {\sigma_h^i}^2$, where $D_h^i = \arg\max_n h_n^i$. The coefficient 7 is due to the fact that the data is reported weekly. For each time-series, we use the log-normal function to model the data after an infection. For a candidate infection time, we find σ_h^i and D_h^i such that the Mean Square Error (MSE) is minimized. We also choose A_h^i such that $d_{D_h^i}^i = h_{D_h^i}^i$. The "individual infection time estimate" is then the candidate infection

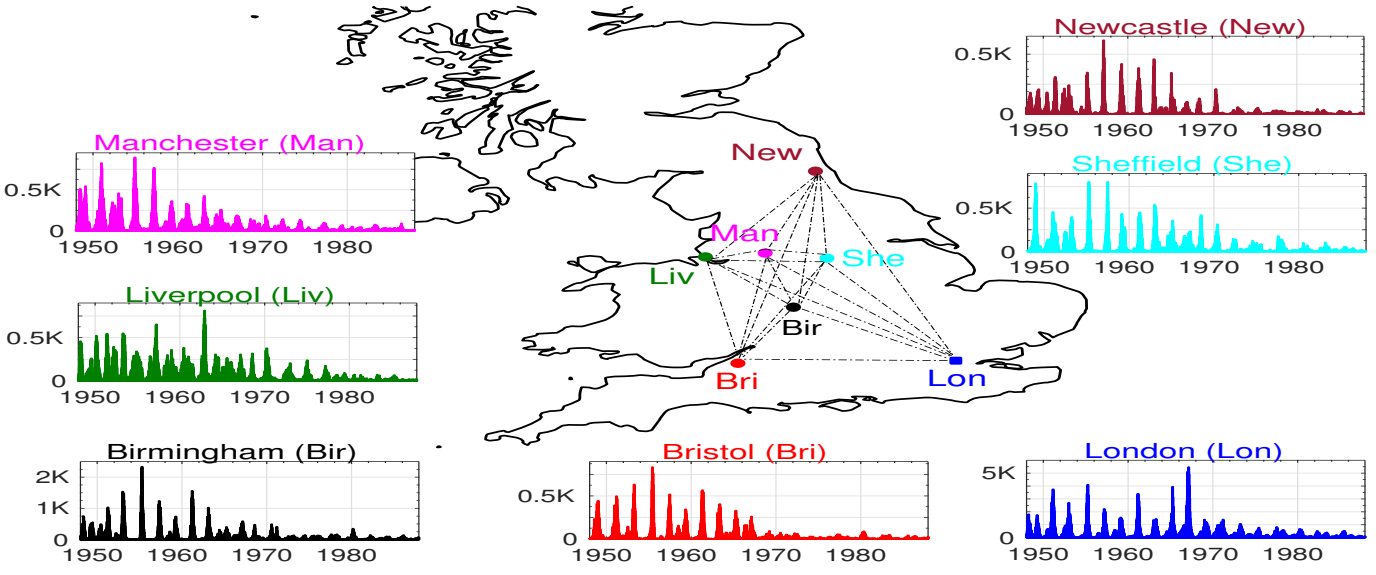


Fig. 1: Weekly Reports of Measles and Chickenpox in 7 Major Cities of England and Wales: London (Lon), Bristol (Bri), Liverpool (Liv), Manchester (Man), Newcastle (New), Birmingham (Bir), Sheffield (She)

time that has minimum MSE after the log-normal fit. We model the residual using a Gaussian distribution. We also assume that the data follows a normal distribution before the individual infection times, i.e.,

$$d_n^i | t_1^i \sim \begin{cases} \mathcal{N}(\mu_1^i, \sigma_1^i) & \text{if } n > t_1^i, \\ \mathcal{N}(h_n^i + \mu_2^i, \sigma_2^i) & \text{if } n \leq t_1^i. \end{cases} \tag{2}$$

In these models, we use empirical variances derived from the previous time-period. Figures 2 to 15 show the quantile-quantile plots for all the seven nodes (cities) in seven successive biennial study periods (1952-1966) to provide support for the Gaussian model.

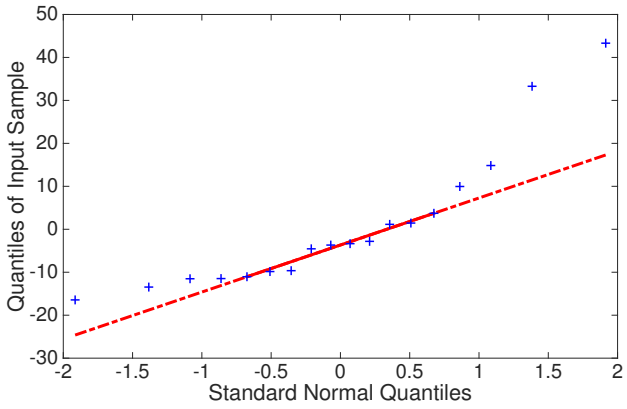
REFERENCES

[1] B. Bolker and B. Grenfell, "Impact of vaccination on the spatial correlation and persistence of measles dynamics," *Proc. Nat. Academy Sci.*, vol. 93, no. 22, pp. 12 648–12 653, 1996.

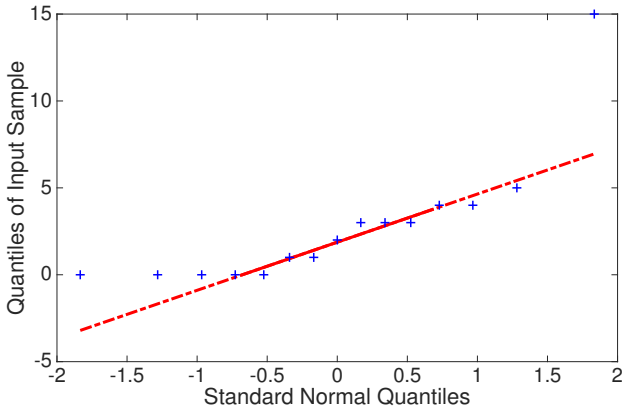
[2] B. Bolker, "Chaos and complexity in measles models: a comparative numerical study," *Math. Med. and Biol.*, vol. 10, no. 2, pp. 83–95, 1993.

[3] W. Wang, Z. Wu, C. Wang, and R. Hu, "Modelling the spreading rate of controlled communicable epidemics through an entropy-based thermodynamic model," *Science China Phys., Mech., and Astro.*, vol. 56, no. 11, pp. 2143–2150, 2013.

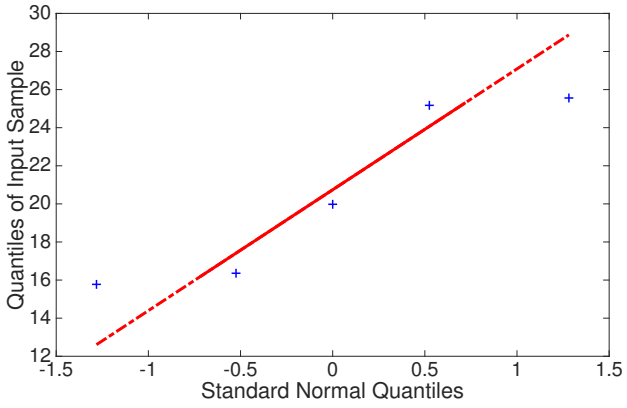
[4] S. Shaghaghian and M. Coates, "Online Bayesian inference of diffusion networks," Dept. Elect. and Comput. Eng., McGill Univ., Montreal, QC, Tech. Rep., 2016.



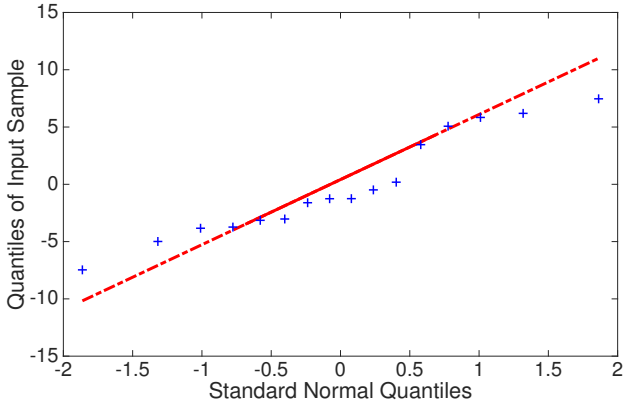
(a) London



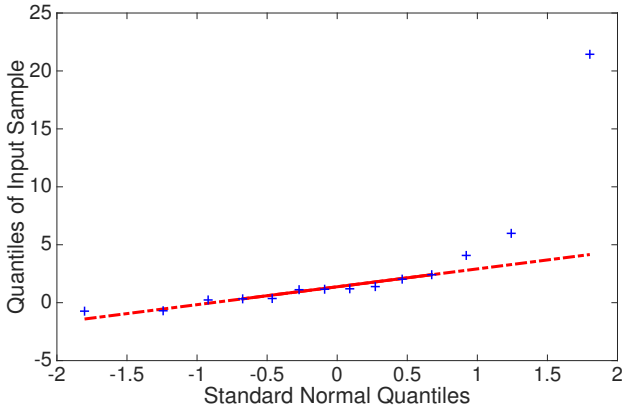
(b) Bristol



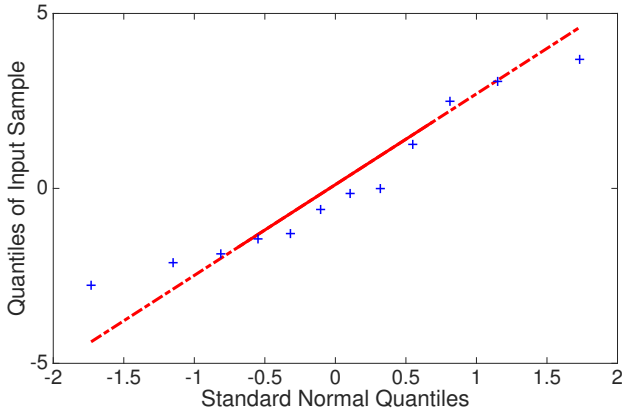
(c) Liverpool



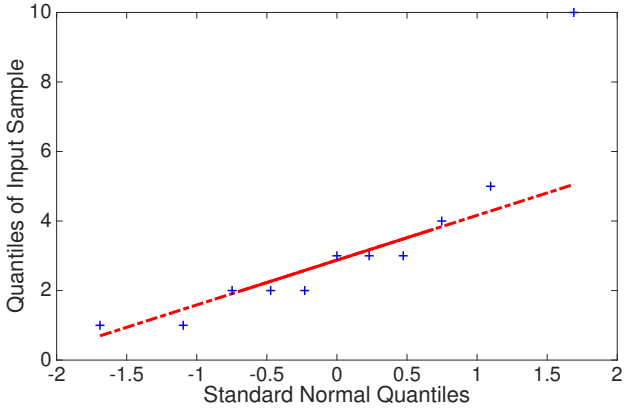
(d) Manchester



(e) Newcastle

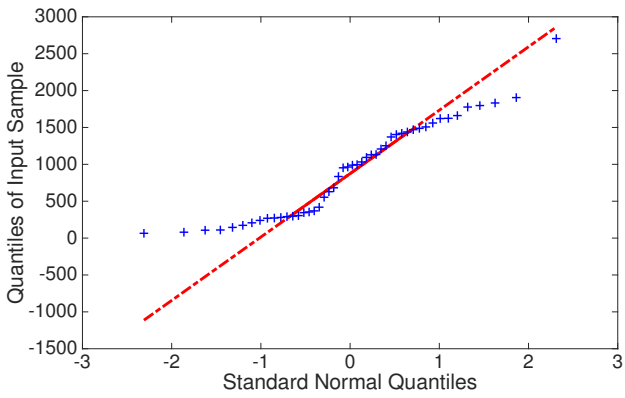


(f) Birmingham

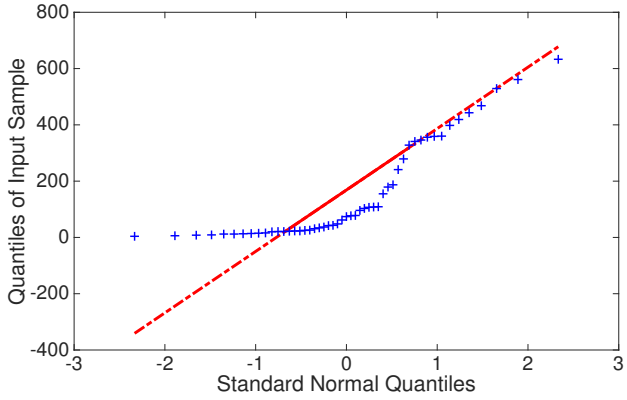


(g) Sheffield

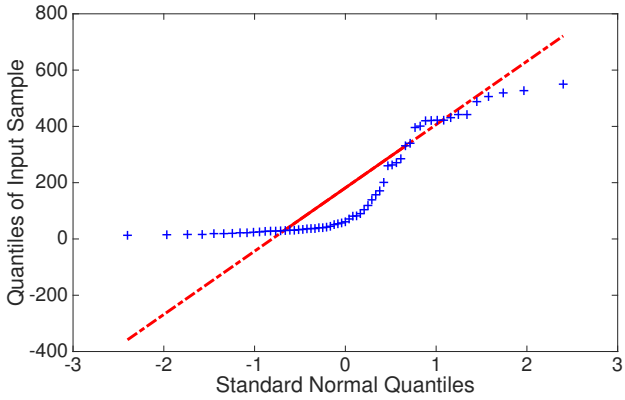
Fig. 2: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1952-1954



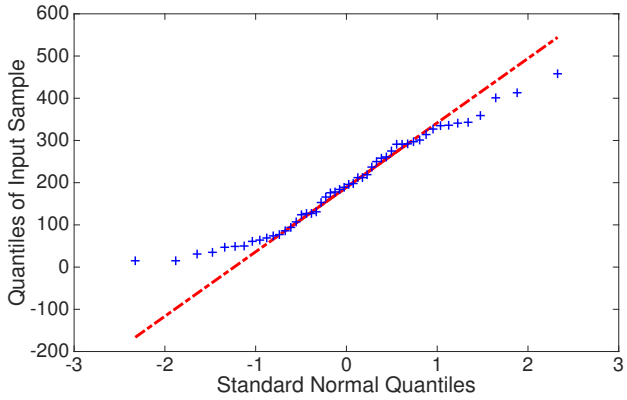
(a) London



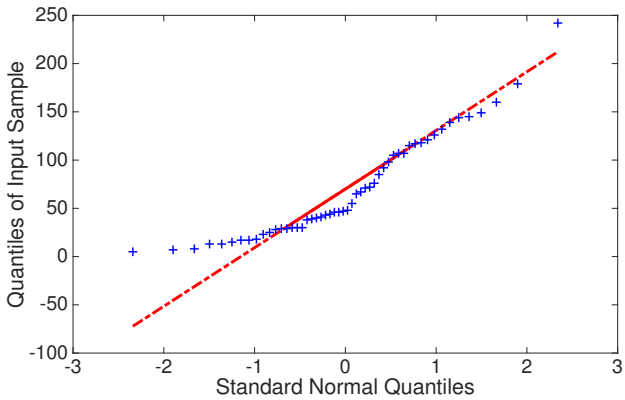
(b) Bristol



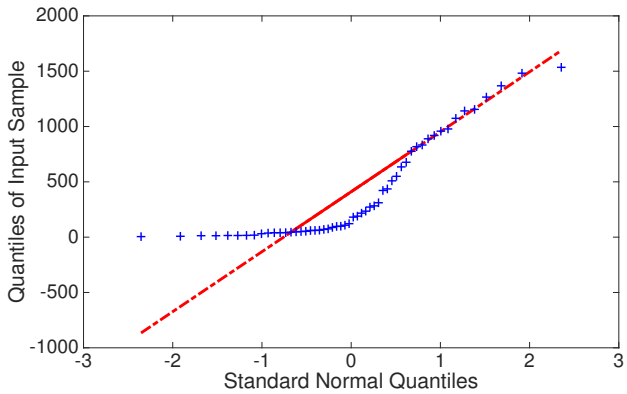
(c) Liverpool



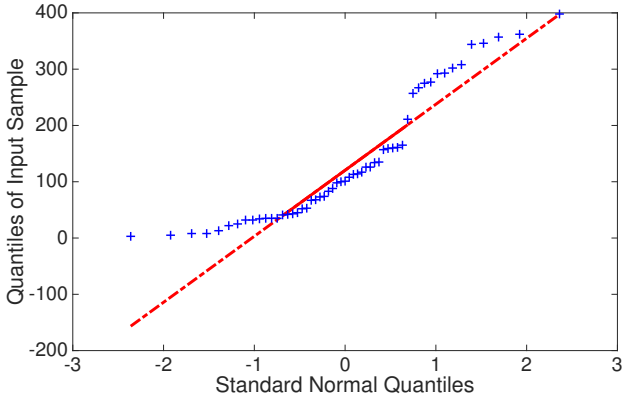
(d) Manchester



(e) Newcastle

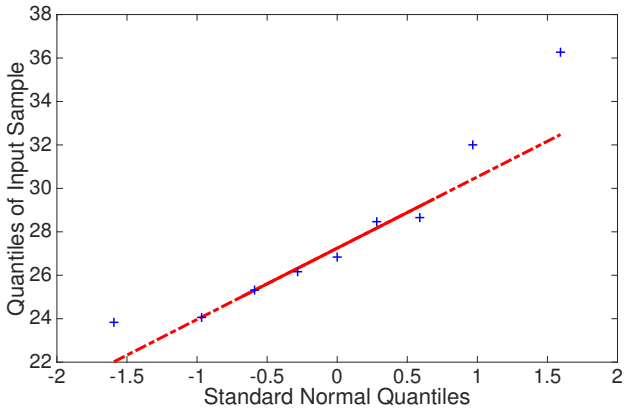


(f) Birmingham

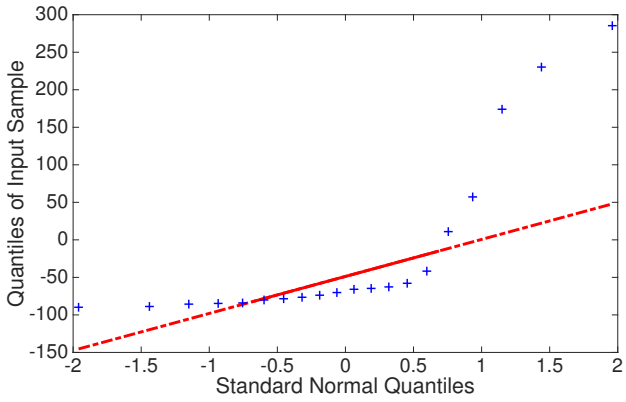


(g) Sheffield

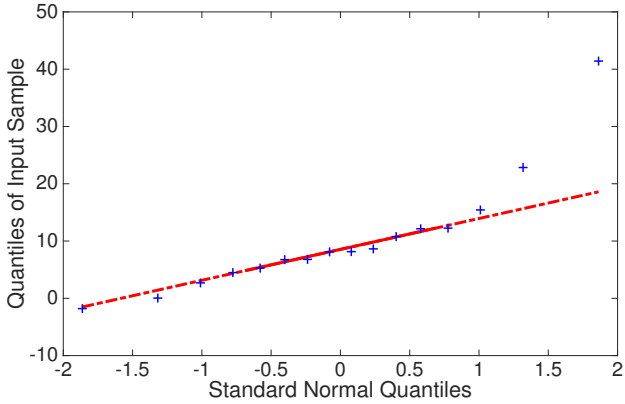
Fig. 3: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1952-1954



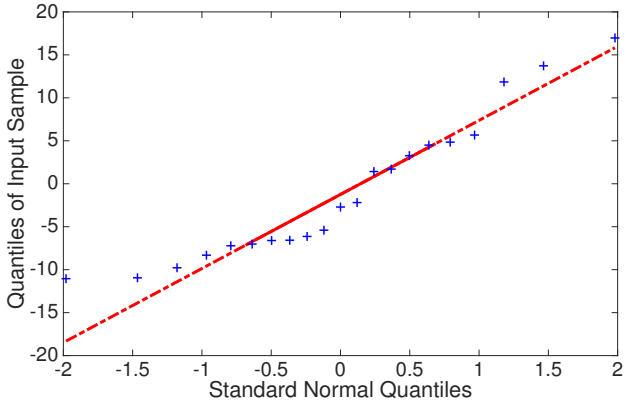
(a) London



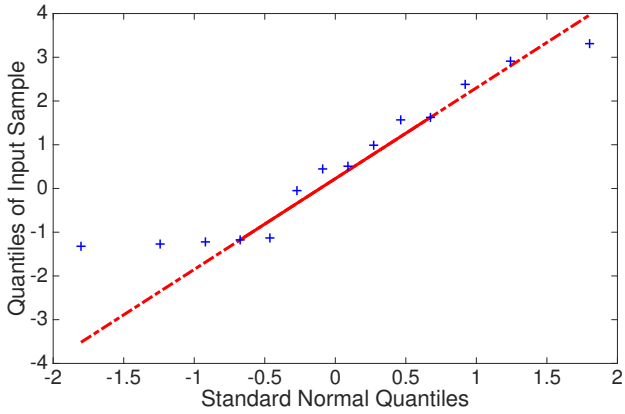
(b) Bristol



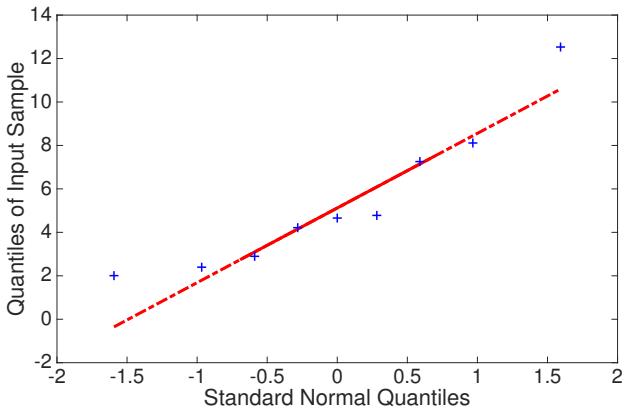
(c) Liverpool



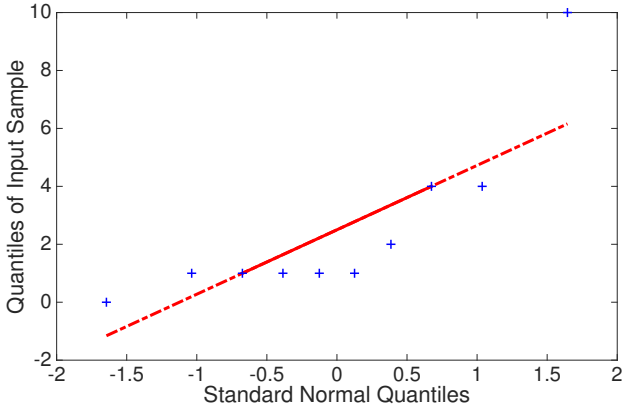
(d) Manchester



(e) Newcastle

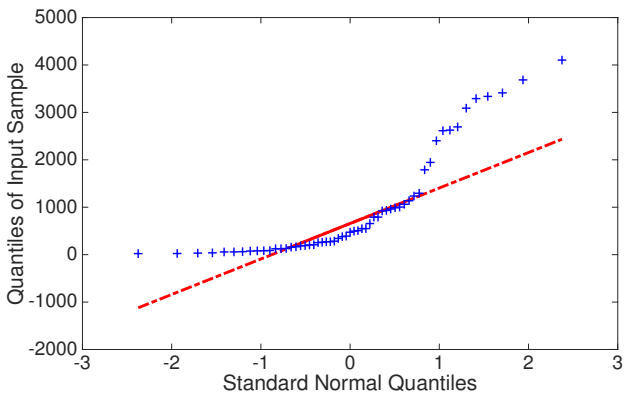


(f) Birmingham

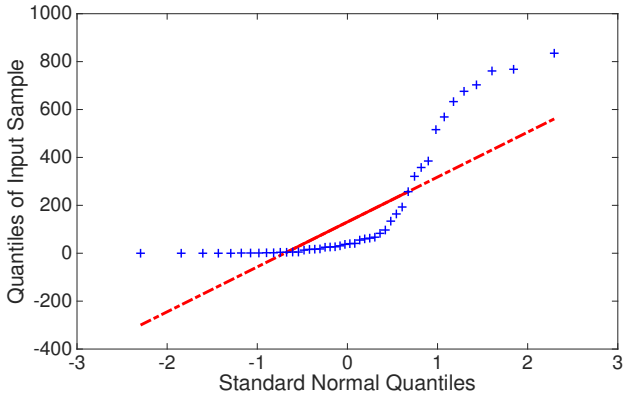


(g) Sheffield

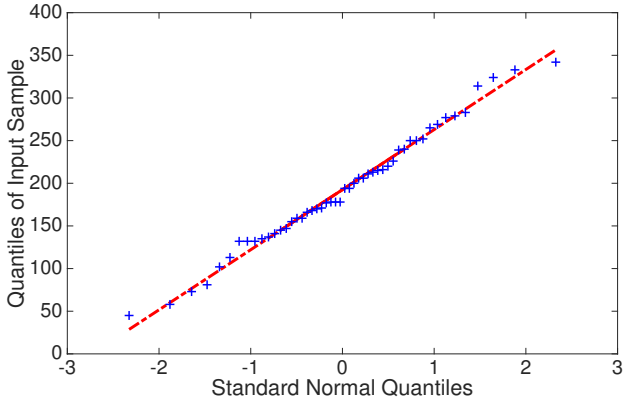
Fig. 4: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1954-1956



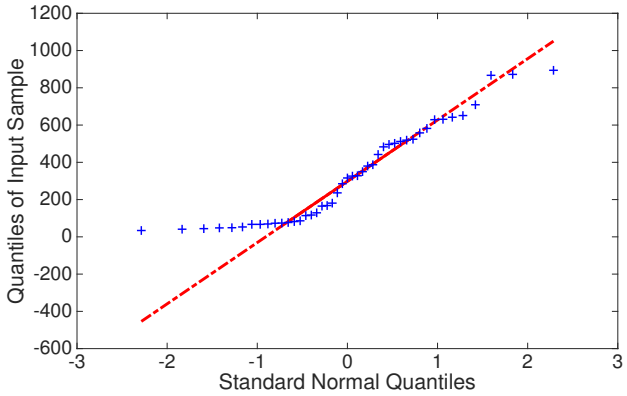
(a) London



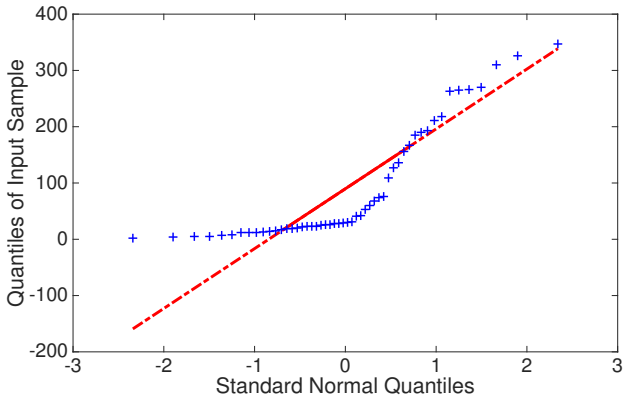
(b) Bristol



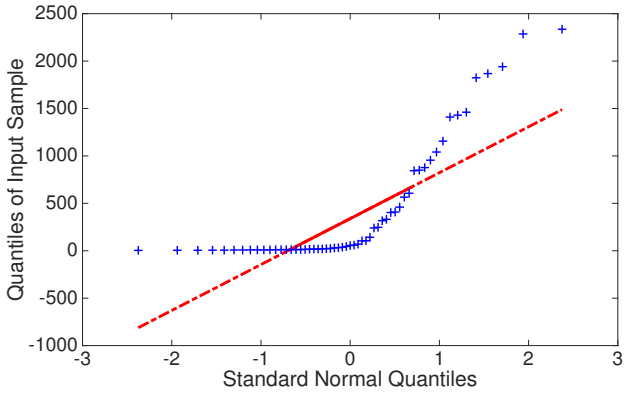
(c) Liverpool



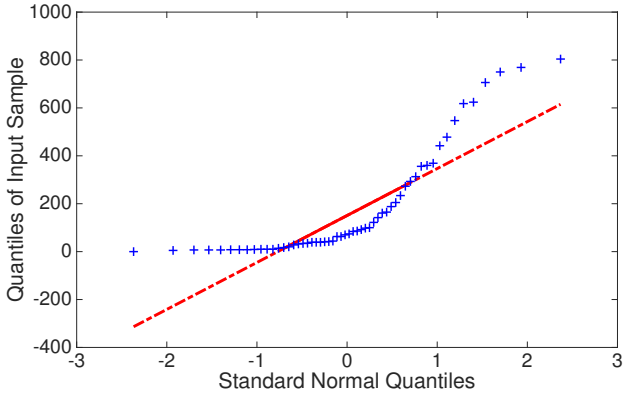
(d) Manchester



(e) Newcastle

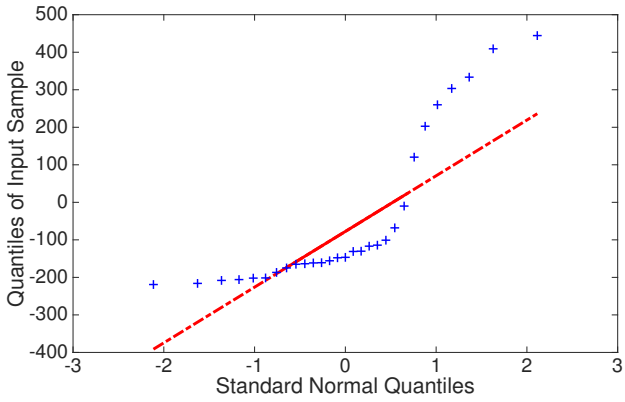


(f) Birmingham

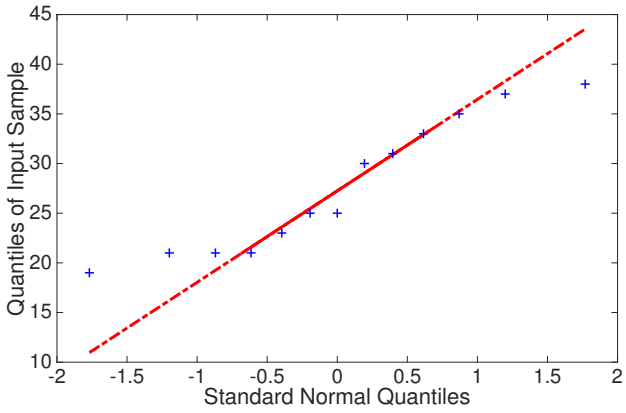


(g) Sheffield

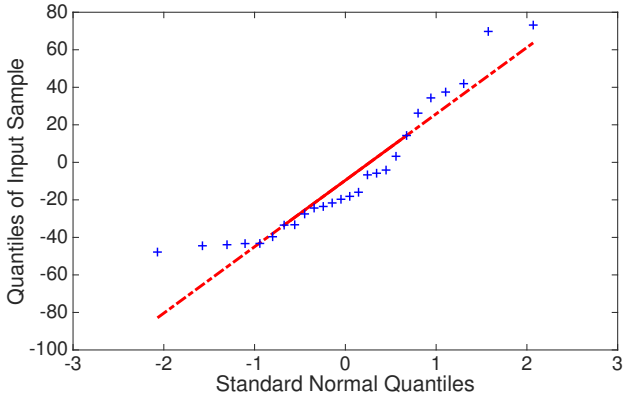
Fig. 5: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1954-1956



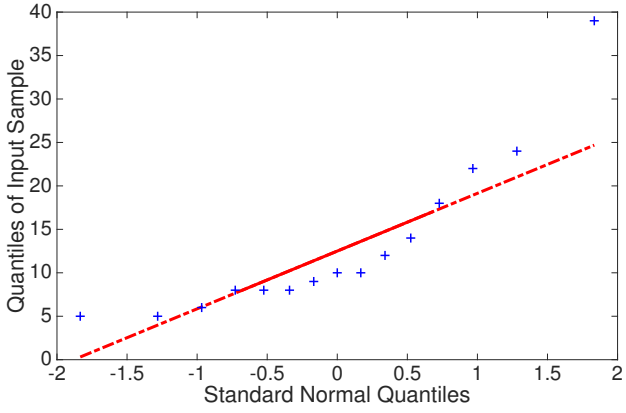
(a) London



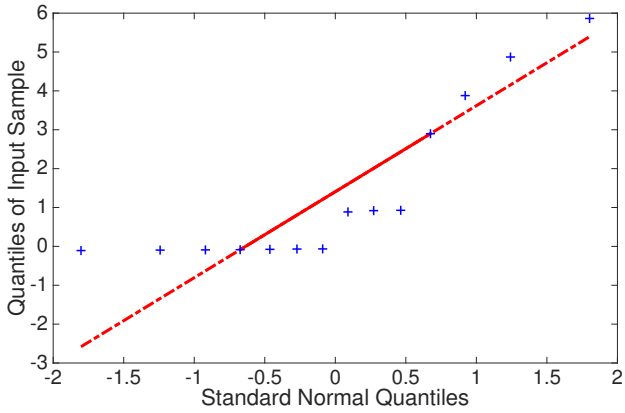
(b) Bristol



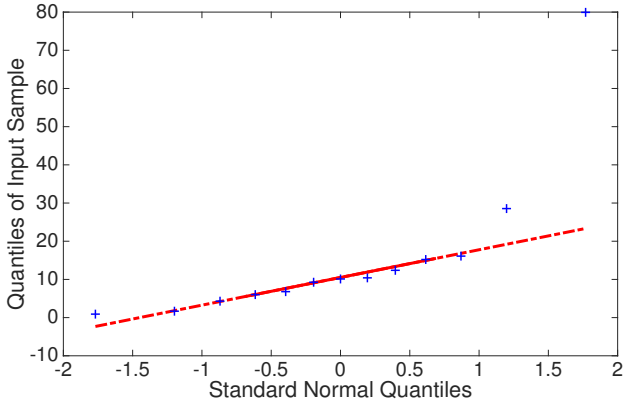
(c) Liverpool



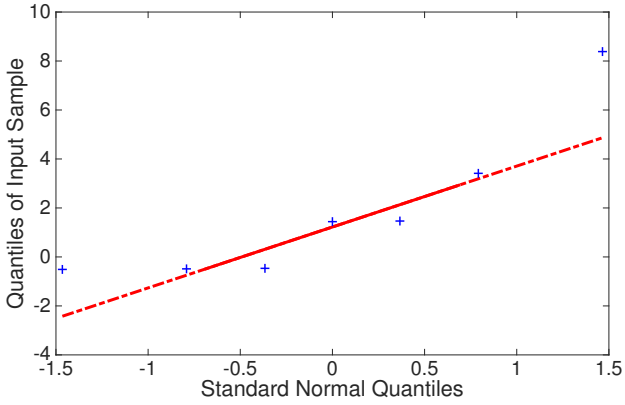
(d) Manchester



(e) Newcastle

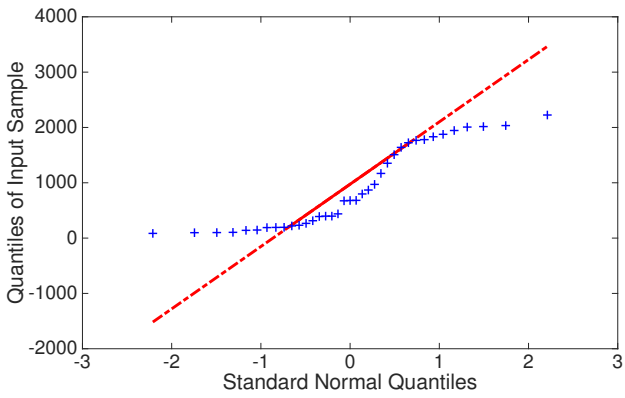


(f) Birmingham

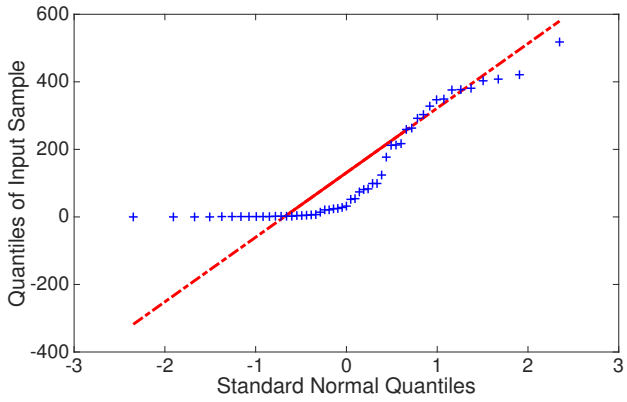


(g) Sheffield

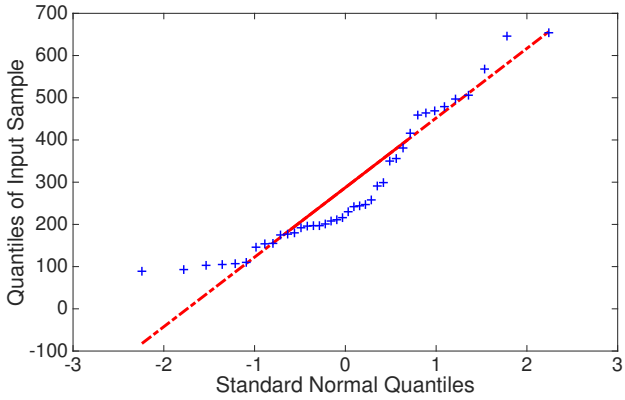
Fig. 6: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1956-1958



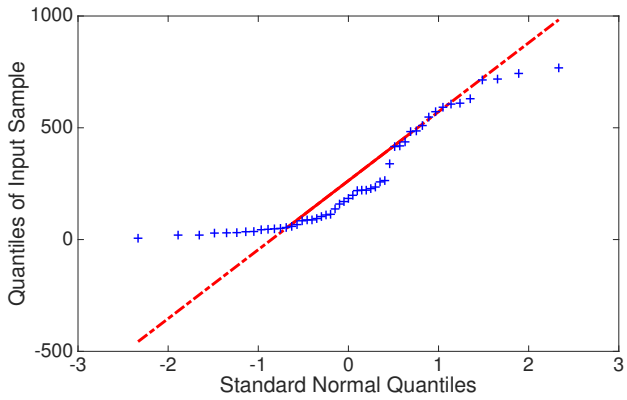
(a) London



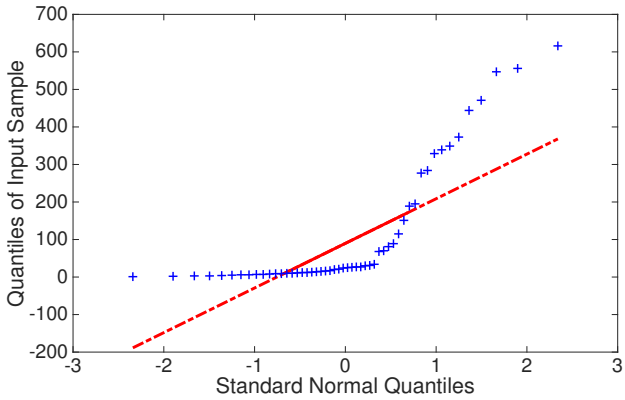
(b) Bristol



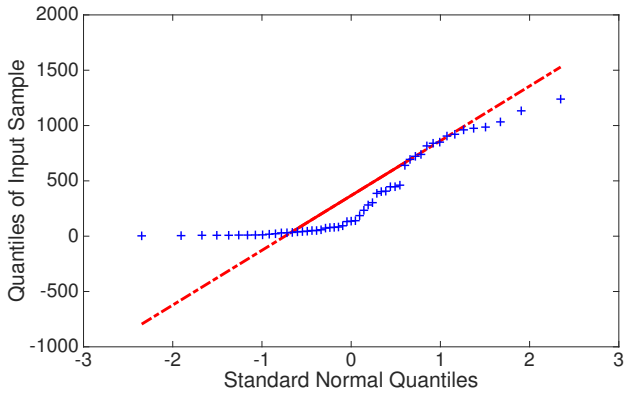
(c) Liverpool



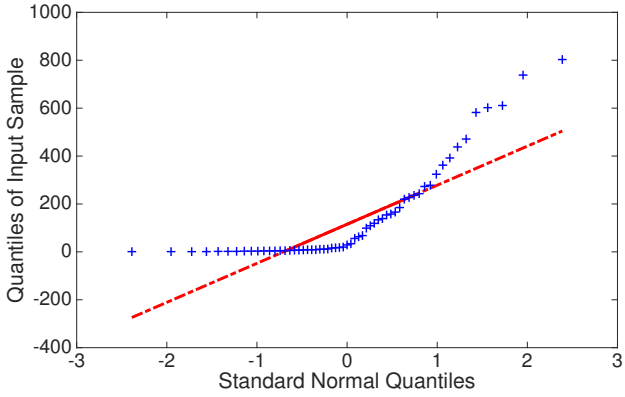
(d) Manchester



(e) Newcastle

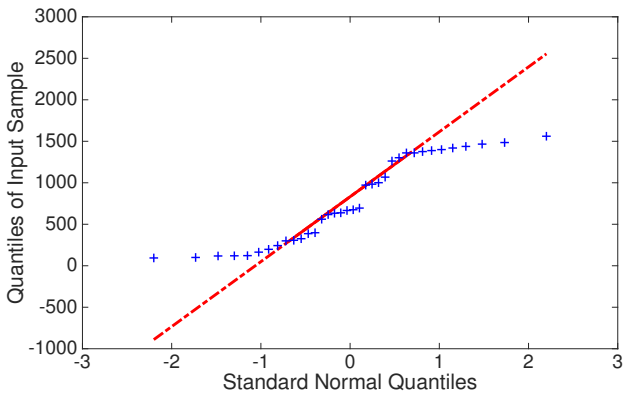


(f) Birmingham

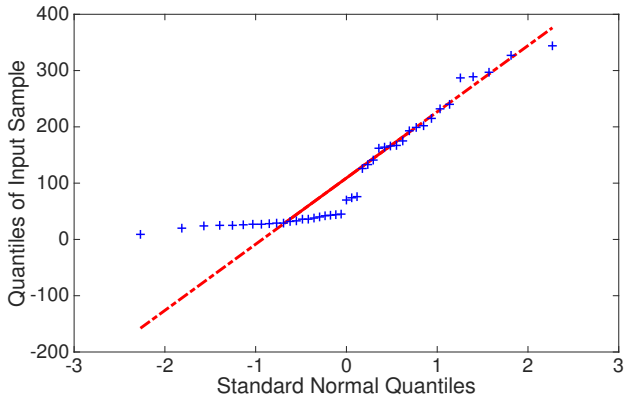


(g) Sheffield

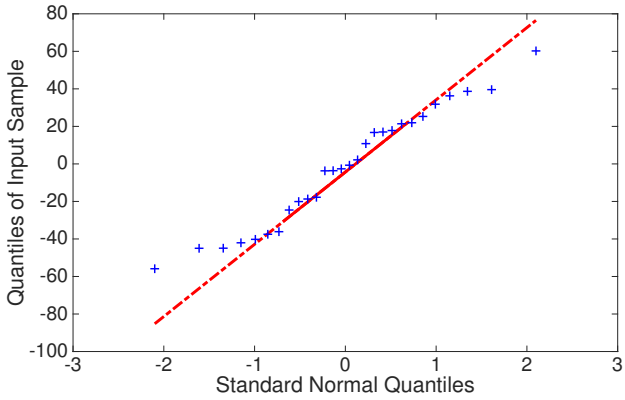
Fig. 7: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1956-1958



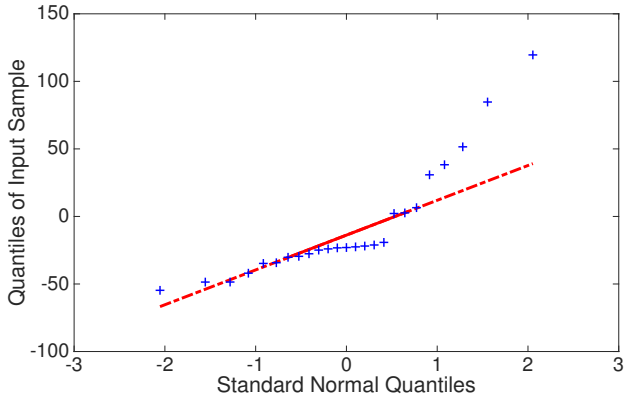
(a) London



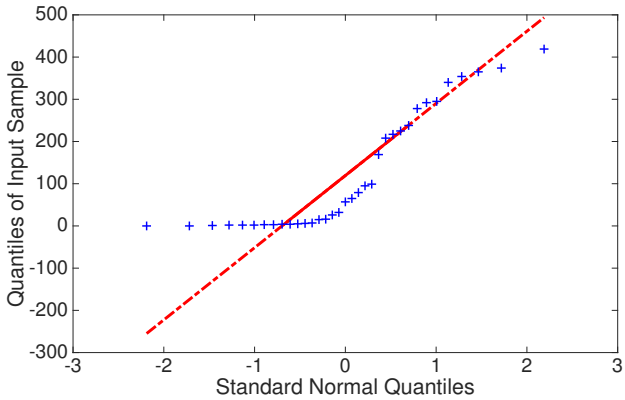
(b) Bristol



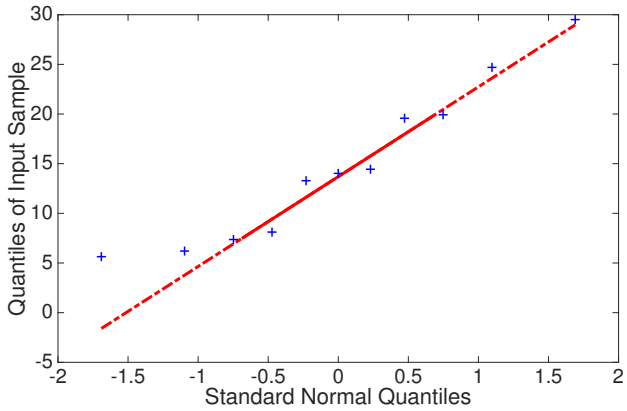
(c) Liverpool



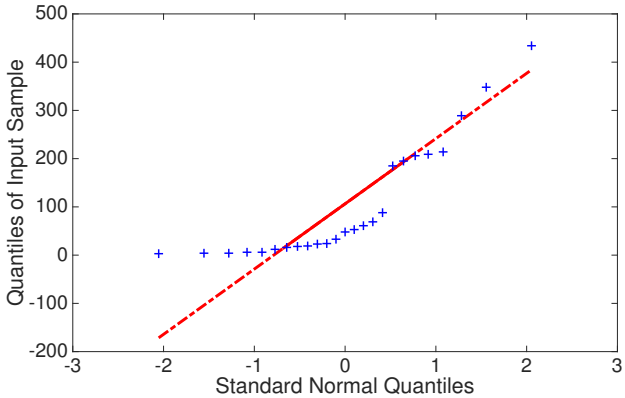
(d) Manchester



(e) Newcastle

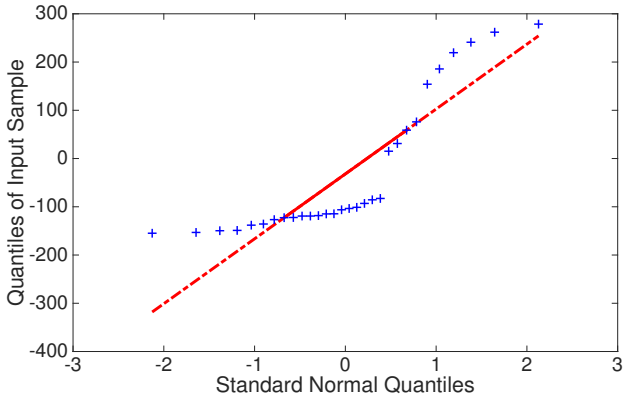


(f) Birmingham

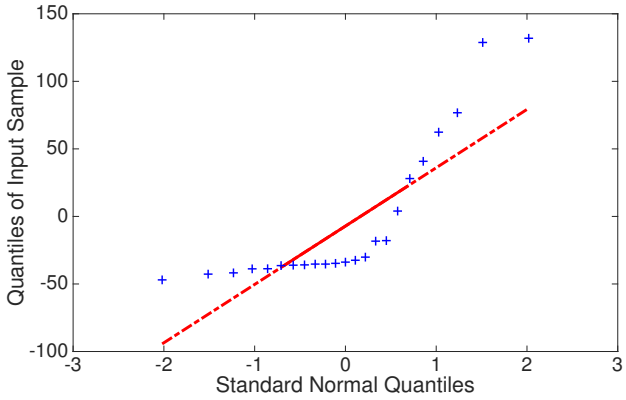


(g) Sheffield

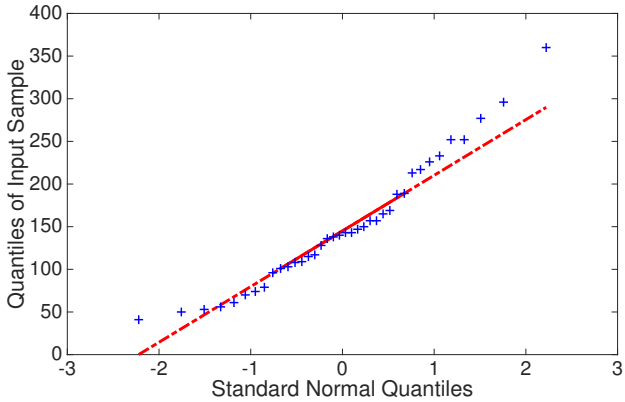
Fig. 8: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1958-1960



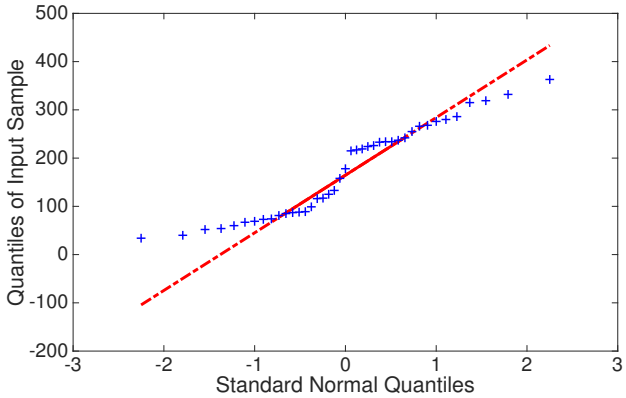
(a) London



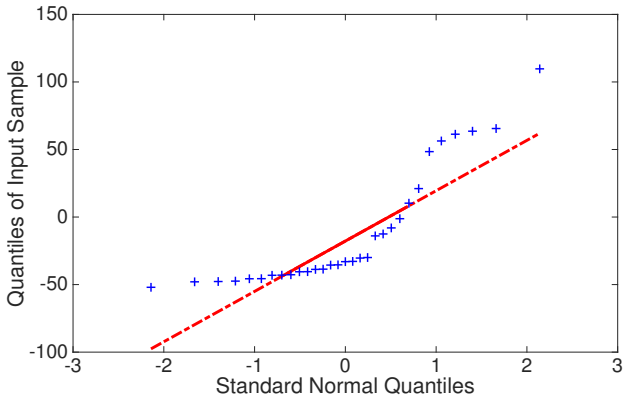
(b) Bristol



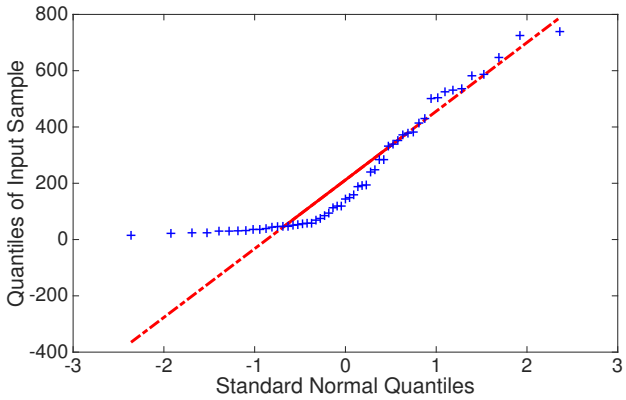
(c) Liverpool



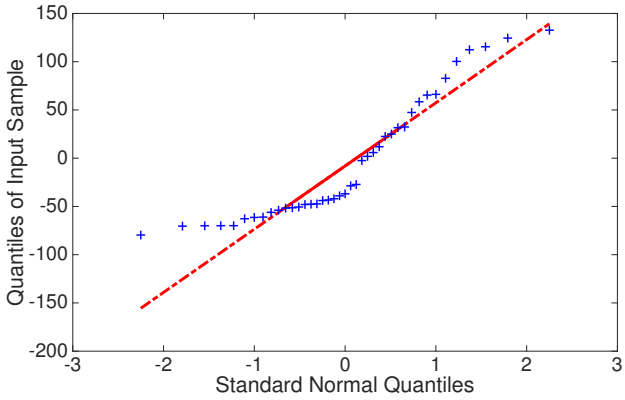
(d) Manchester



(e) Newcastle

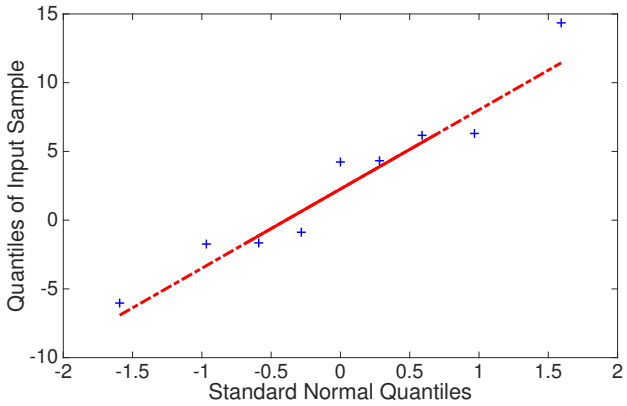


(f) Birmingham

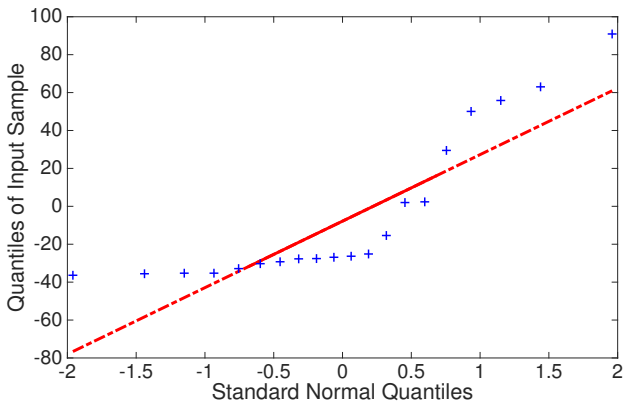


(g) Sheffield

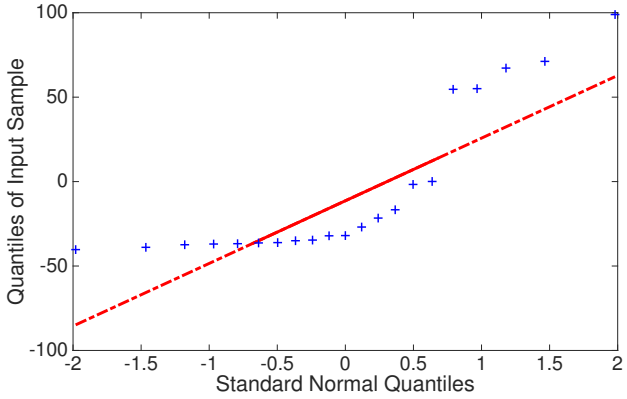
Fig. 9: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1958-1960



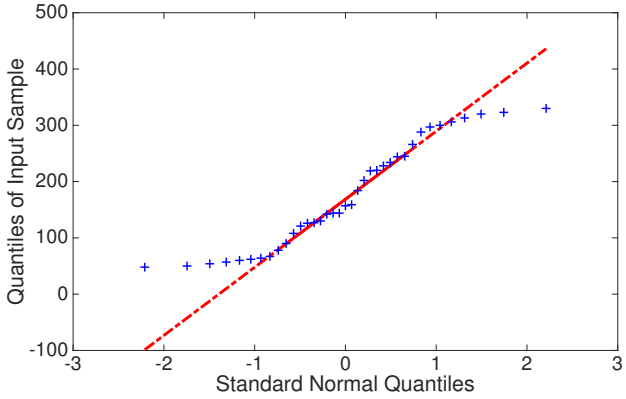
(a) London



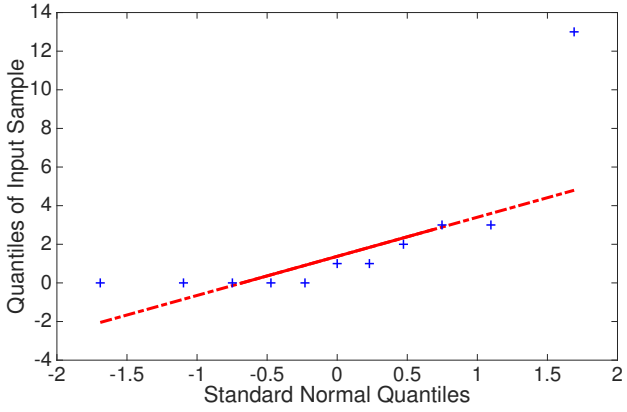
(b) Bristol



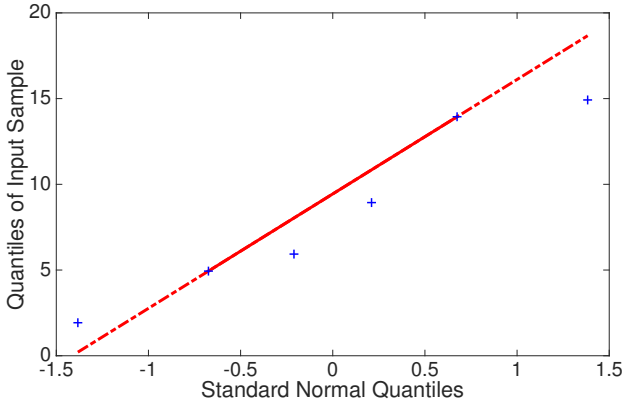
(c) Liverpool



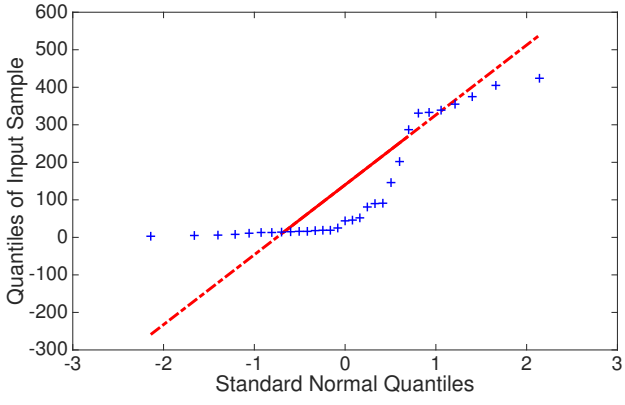
(d) Manchester



(e) Newcastle

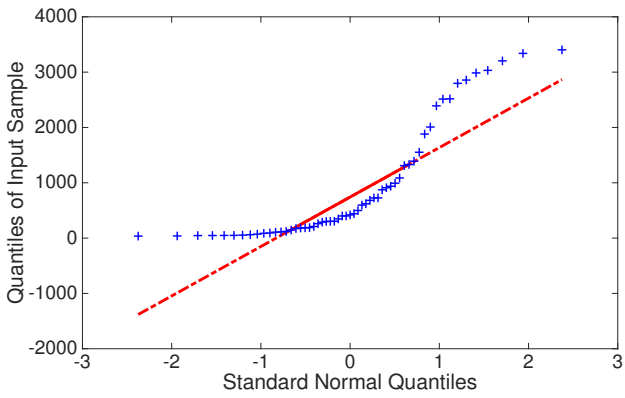


(f) Birmingham

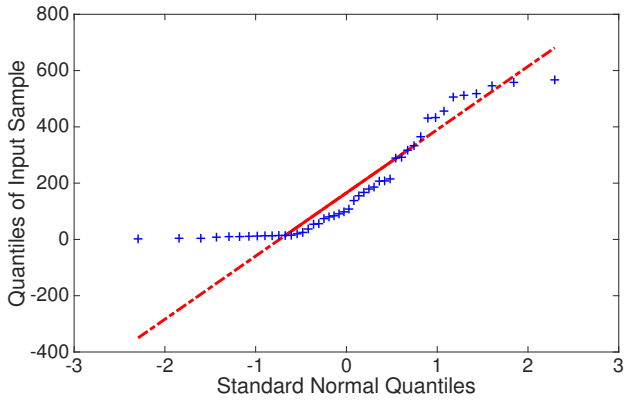


(g) Sheffield

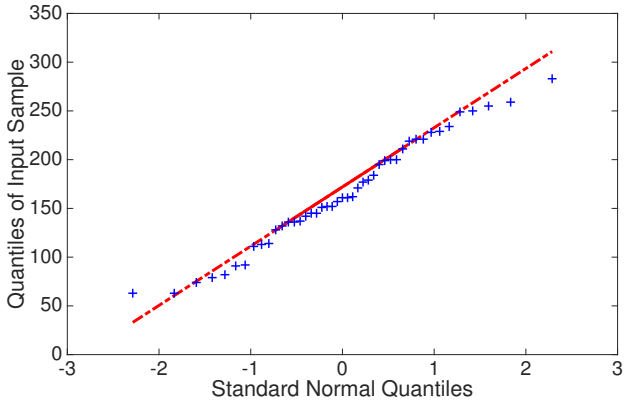
Fig. 10: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1960-1962



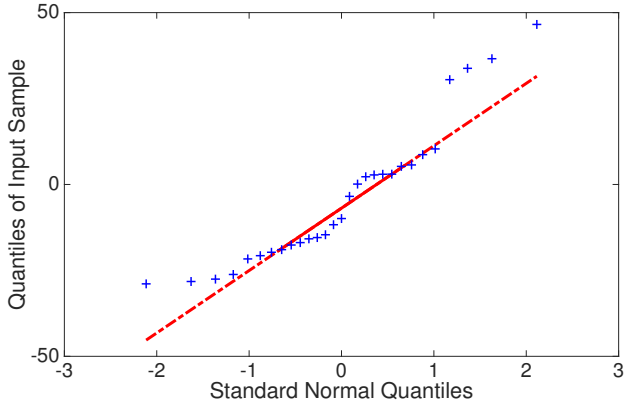
(a) London



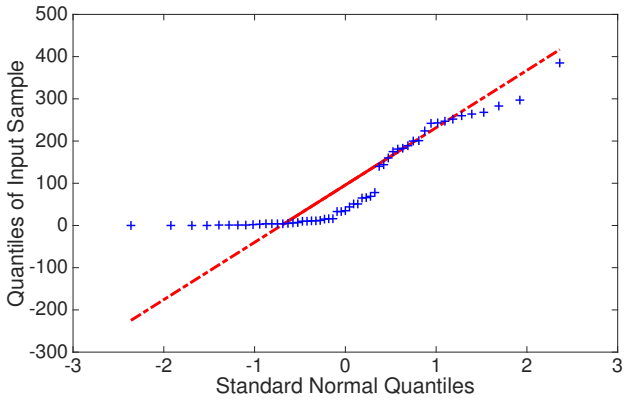
(b) Bristol



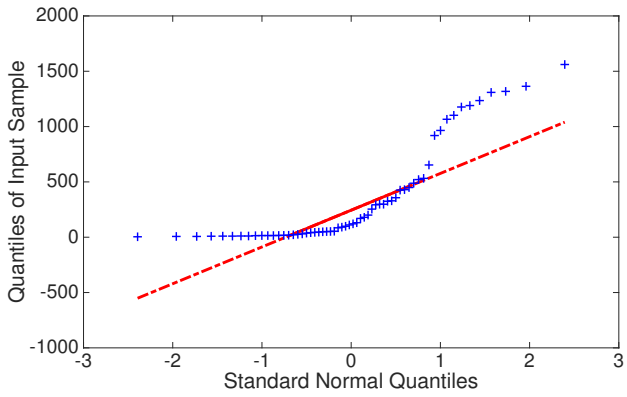
(c) Liverpool



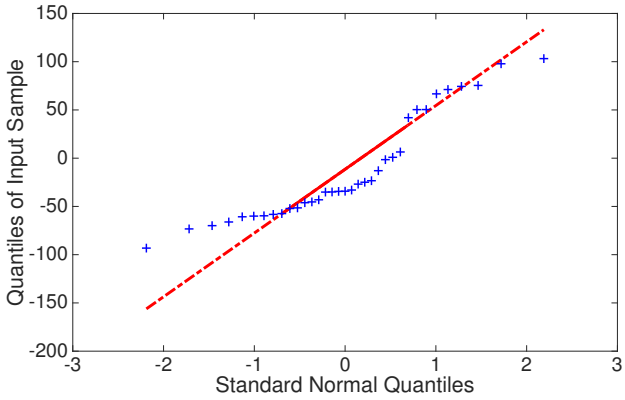
(d) Manchester



(e) Newcastle

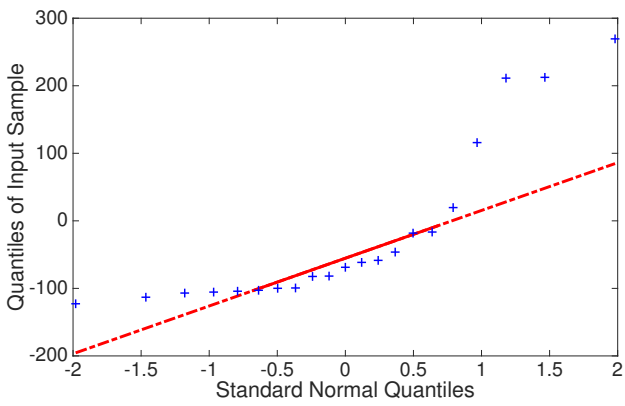


(f) Birmingham

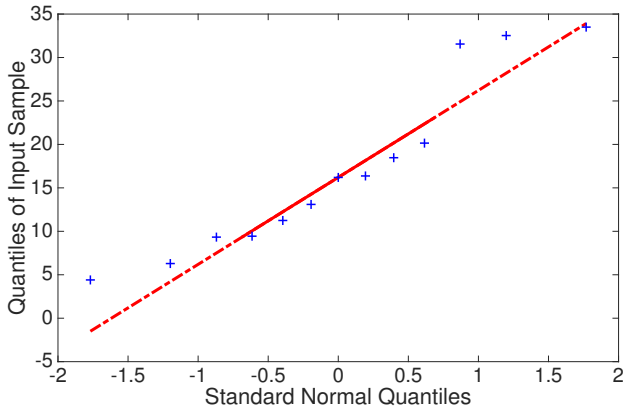


(g) Sheffield

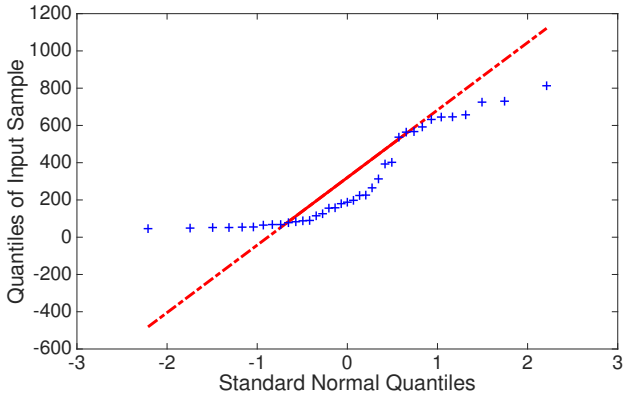
Fig. 11: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1960-1962



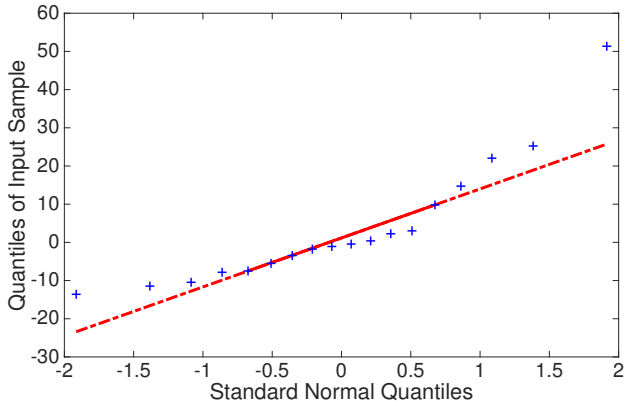
(a) London



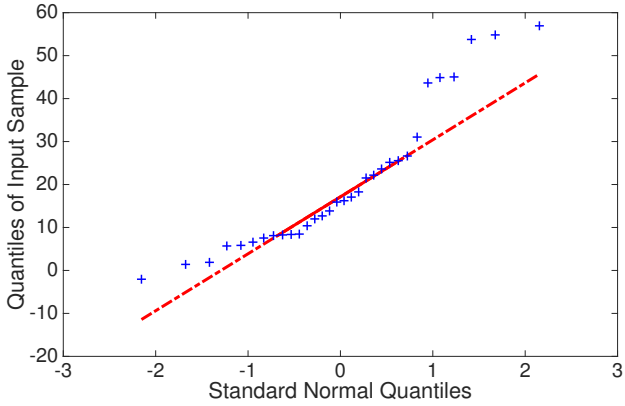
(b) Bristol



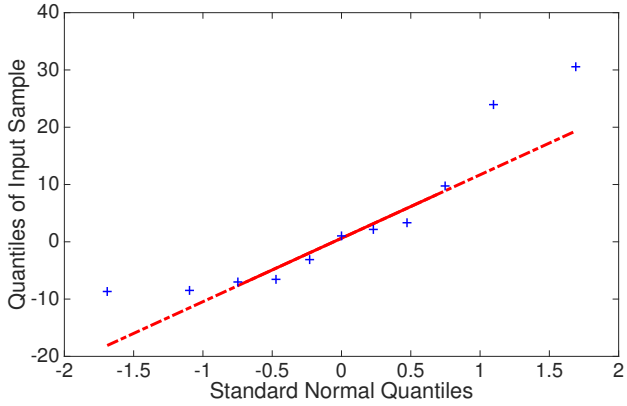
(c) Liverpool



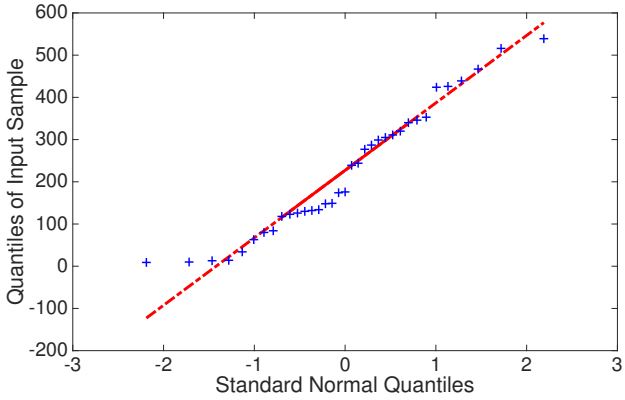
(d) Manchester



(e) Newcastle

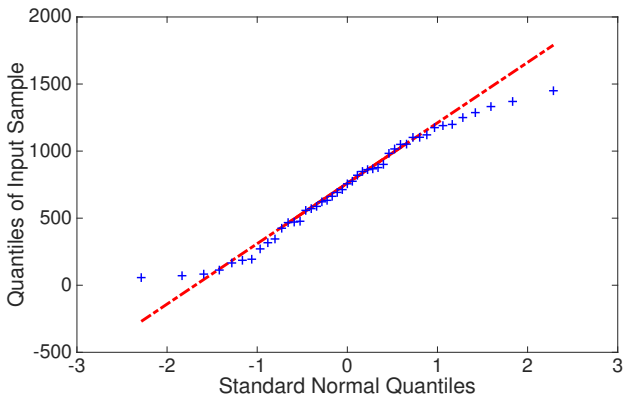


(f) Birmingham

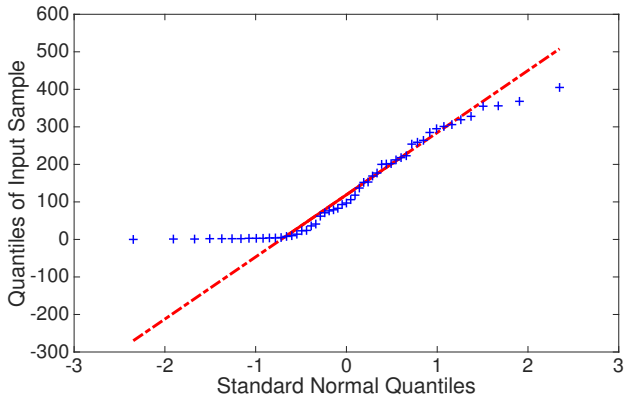


(g) Sheffield

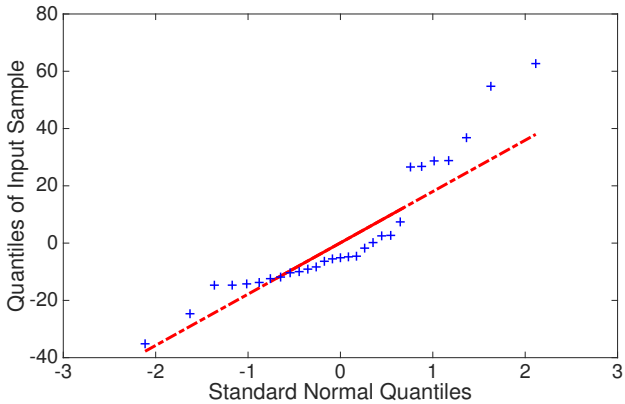
Fig. 12: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1962-1964



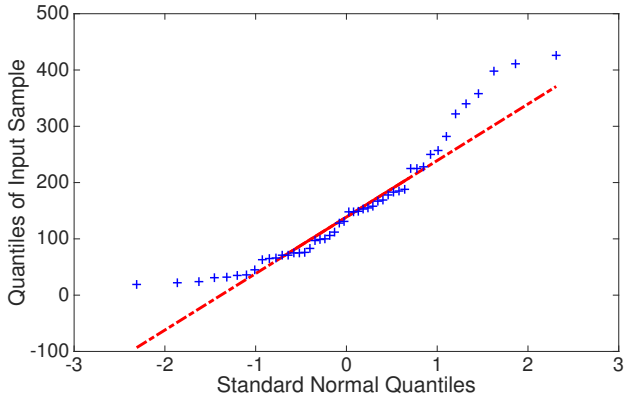
(a) London



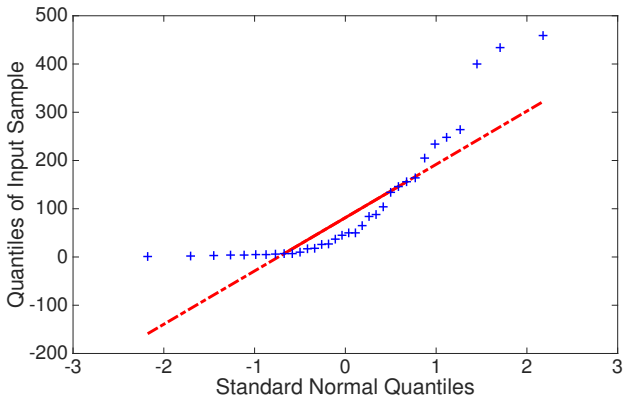
(b) Bristol



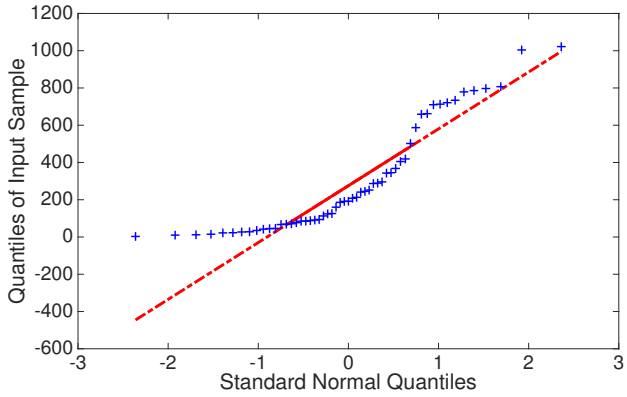
(c) Liverpool



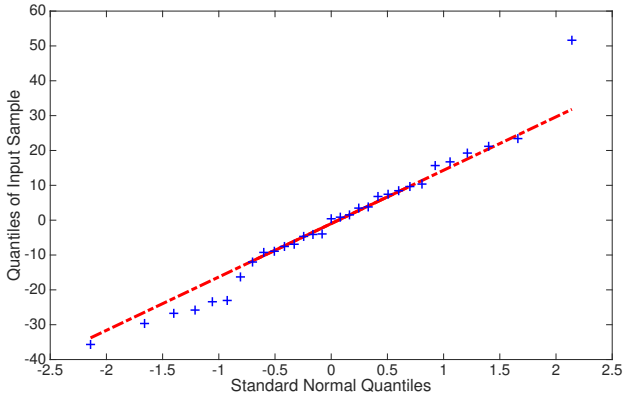
(d) Manchester



(e) Newcastle

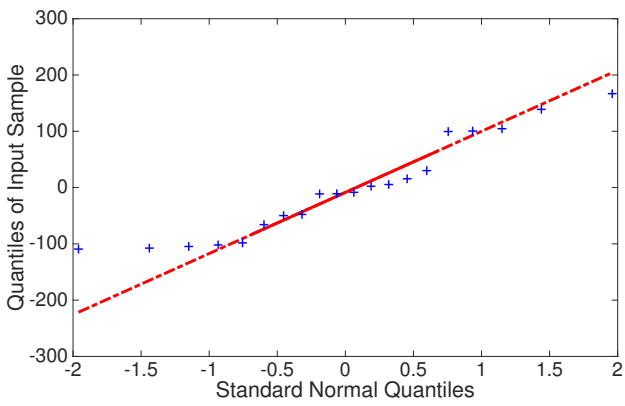


(f) Birmingham

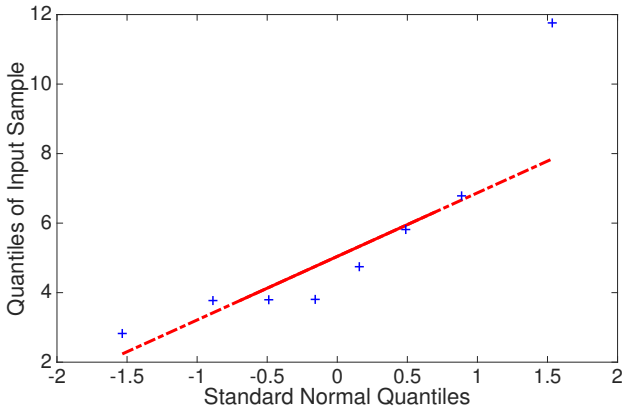


(g) Sheffield

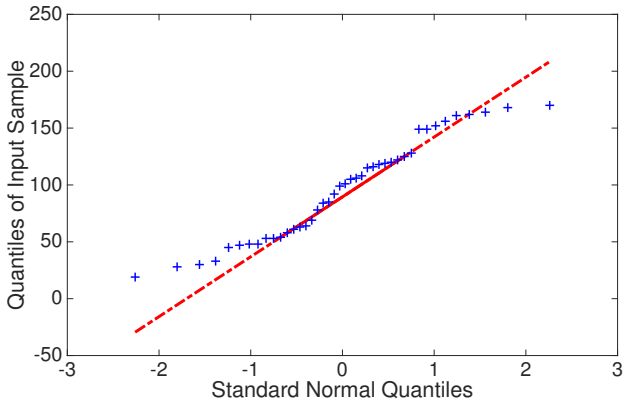
Fig. 13: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1962-1964



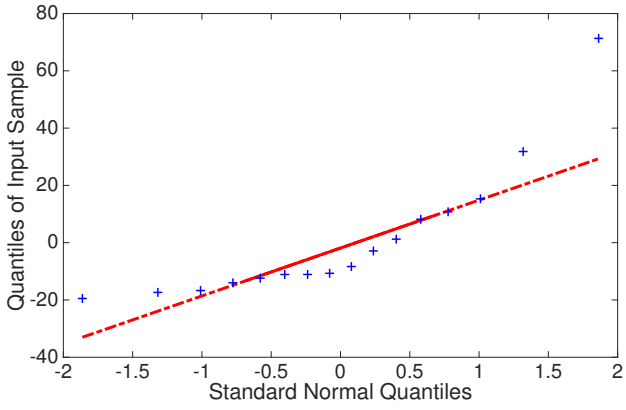
(a) London



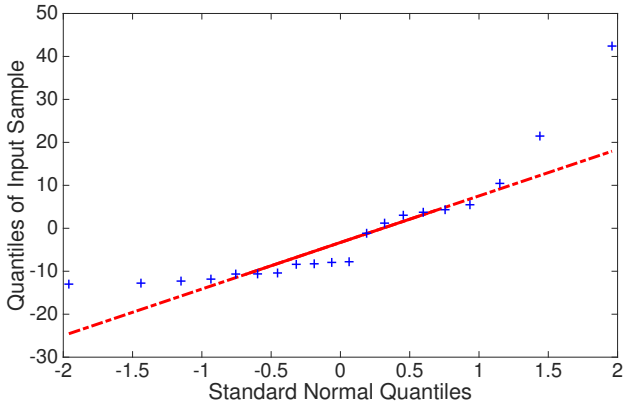
(b) Bristol



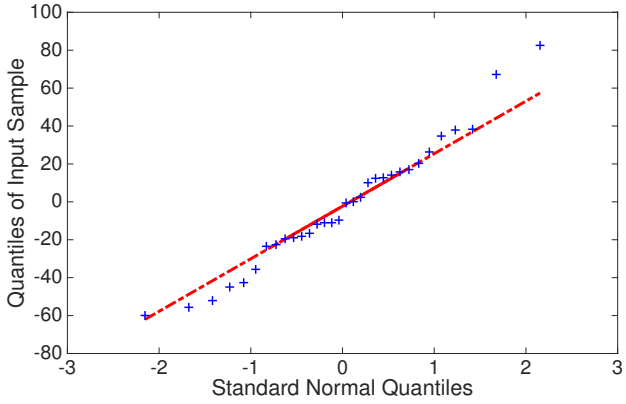
(c) Liverpool



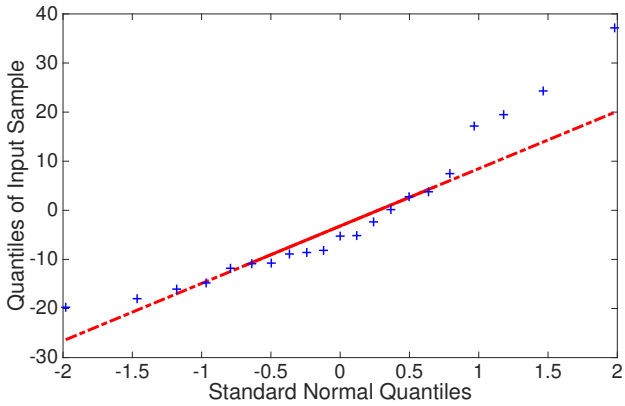
(d) Manchester



(e) Newcastle

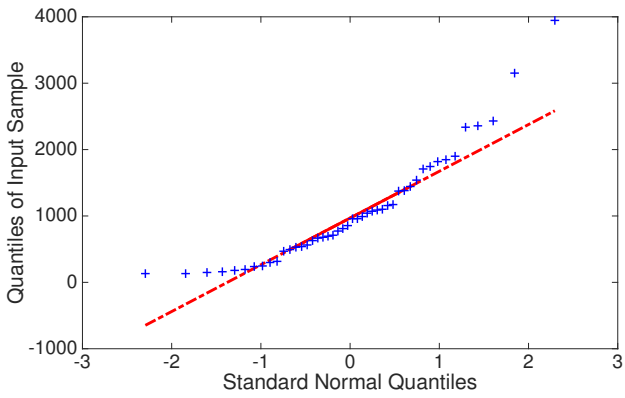


(f) Birmingham

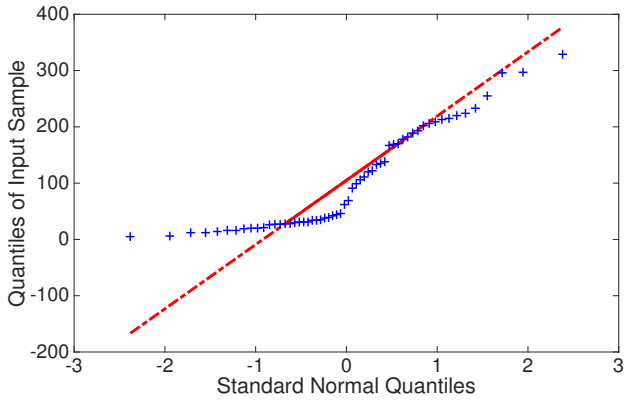


(g) Sheffield

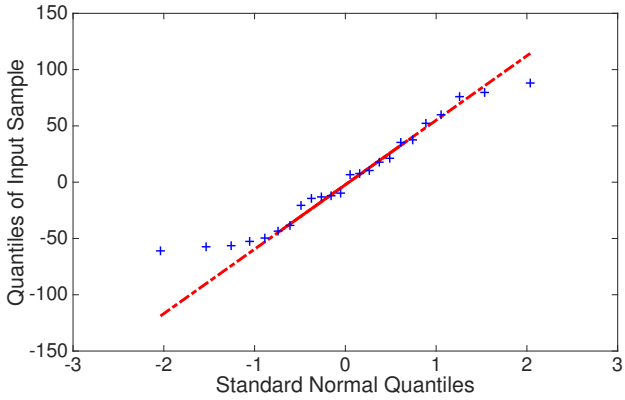
Fig. 14: Quantile-quantile (qq) plot of the reported data before the individual infection time versus standard normal distribution in 1964-1966



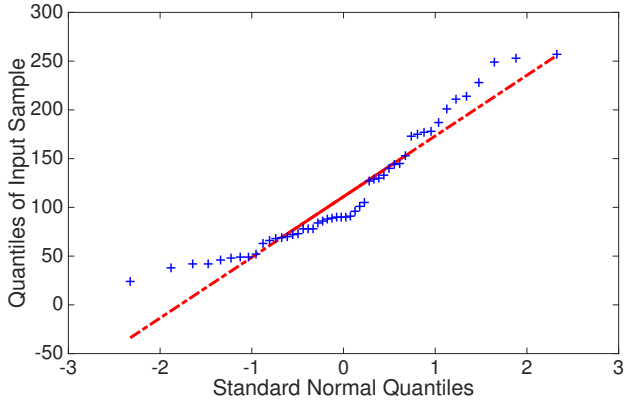
(a) London



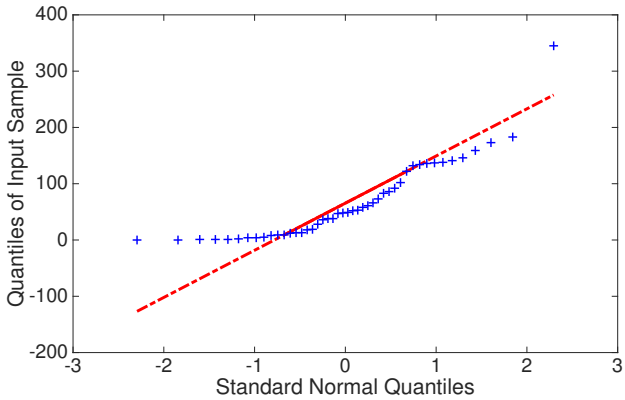
(b) Bristol



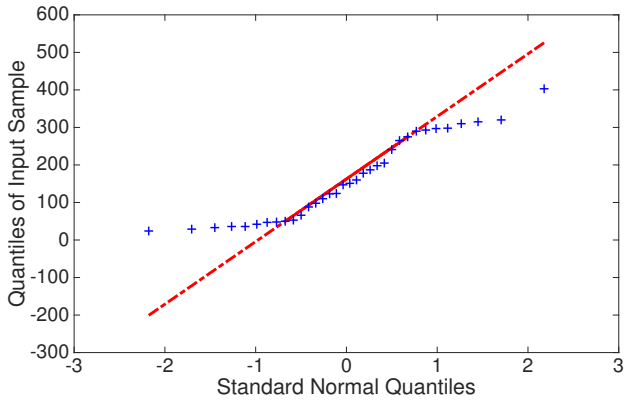
(c) Liverpool



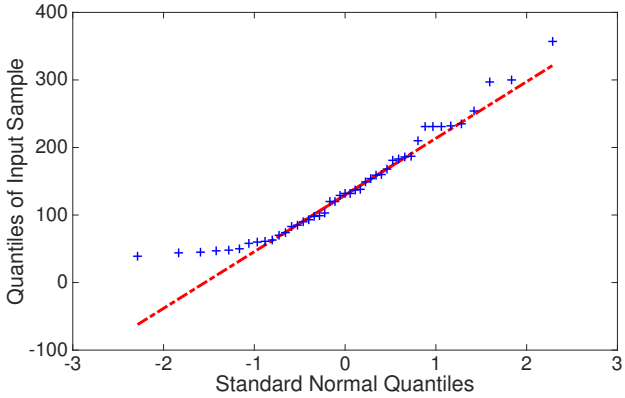
(d) Manchester



(e) Newcastle



(f) Birmingham



(g) Sheffield

Fig. 15: Quantile-quantile (qq) plot of the residual noise after the individual infection time versus standard normal distribution in 1964-1966