

THE ENVIRONMENTAL ACCIDENT AT 'SCHWEIZERHALLE' AND RESPIRATORY DISEASES IN CHILDREN: A TIME SERIES ANALYSIS

ULRICH HELFENSTEIN

*Biostatistisches Zentrum, Institut für sozial- und Präventivmedizin, Universität Zürich, Sumatrastrasse 30,
CH-8006 Zurich, Switzerland*

URSULA ACKERMANN-LIEBRICH AND CHARLOTTE BRAUN-FAHRLÄNDER

Abteilung für Sozial- und Präventivmedizin, Universität Basel, Steinengraben 49, CH-4051 Basle, Switzerland

AND

HANS URS WANNER

Institut für Hygiene und Arbeitsphysiologie, ETH Zürich, ETH-Zentrum, CH-8092 Zurich, Switzerland

SUMMARY

During an investigation concerned with the relationship between air pollution and respiratory diseases in children, the 'Schweizerhalle' accident occurred when unknown amounts of pollutants were discharged into the environment. In that investigation, two series of medical data were collected during one year: (a) The daily relative number of preschool children, exhibiting diseases of the respiratory tract, who either came to the outpatients' clinic of the Children's Hospital or were reported by paediatricians in Basle; (b) The daily number of respiratory symptoms per child, observed in a group of randomly selected preschool children. The purpose of the present time series analysis is the assessment of possible change in these series after the environmental accident. The nature of the change is studied by complementary approaches. First, a forecast arising from models identified in the preaccident period is compared with the actual data. Thereafter, intervention models which adequately and parsimoniously represent the change are identified. Finally, an identification of a change-point is performed.

INTRODUCTION

Between 1 November 1985 and 23 November 1986 daily medical and environmental data were collected to study the relationship between respiratory diseases in children and air pollution.¹ Results from this investigation analysed by means of transfer function models have been presented in a previous case study.² During the time of this study the environmental accident at 'Schweizerhalle' occurred.

On 1 November 1986 a Sandoz storehouse containing about 1300 tons of chemical substances (mainly agrochemicals) burned down in 'Schweizerhalle', located approximately 3 miles outside Basle, a city with approximately 200,000 inhabitants. This fire was discovered at 0:30 h. At 3:00 h the firebrigade decided to use water to extinguish it. This water was discharged into the river Rhine causing what has been called 'one of the worst chemical spills ever'.³ Its effect on the Rhine has been studied in depth.⁴ When the water first came into contact with the fire, a foul smelling

0277-6715/91/101481-12\$06.00

© 1991 by John Wiley & Sons, Ltd.

Received August 1990

Revised February 1991

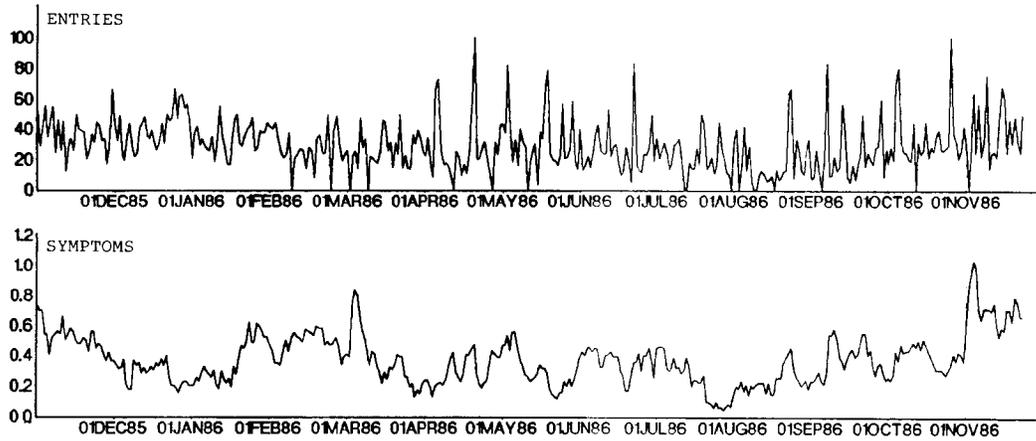


Figure 1. Upper curve: Relative number of entries with respiratory diseases in per cent (ENTRIES). Second curve: Number of symptoms per child (SYMPTOMS)

cloud developed and was carried into the city of Basle. At 4:30 the smell in the city was so intense that the authorities decided to warn the public using sirens, police car megaphones and public broadcasts, which informed people that they should keep windows closed and remain indoors. All traffic was stopped at the city boundary and only at 7:00 h were the warnings withdrawn. Authorities assured the public after 7:00 h that no health hazard existed. But, after many people experienced symptoms and additionally when dead fish appeared in the Rhine, public pressure demanded investigation of possible health effects. In addition to studies specially set up for this purpose, it seems appropriate to analyse ongoing studies to investigate whether adverse health effects could be discovered on or just after the day of the accident.

The study mentioned above was the only one in progress at the date of the accident. In that study two daily series of medical data were collected:

1. The daily number of preschool children with 'respiratory diseases' divided by the daily number of all children who either came to the outpatients' clinic at the Children's Hospital or were reported by paediatricians in Basle. The following were defined to be 'respiratory diseases': rhinopharyngitis, bronchitis, sinusitis, pneumonia, asthma, pseudocroup, otitis and angina. This series is termed 'ENTRIES'.
2. From the birth registry of Basle a random sample of children (aged 0 to 4 years) had been selected. Of the 773 selected families, 575 (74 per cent) participated in the study. About 1/12 of the sample was contacted every month and one child from each family participated in the study for 6 weeks. Parents recorded respiratory symptoms on daily diary forms. At the time of the accident the last group of children were participating in the study. Details of methodology have been reported previously.¹ For the present analysis the daily series of 'number of respiratory symptoms per child' was available. The following respiratory symptoms were recorded: cough, runny or stuffy nose, sore throat, earache and fever (more than 38°C). This series will be termed 'SYMPTOMS'.

Figure 1 shows these two series. A vertical bar indicates the 1 November 1986 (a Saturday). After this date, the series SYMPTOMS exhibits an increase. The corresponding period for the preceding year (November 1985) shows no corresponding behaviour.

The purpose of the present investigation is the assessment of a possible change in these series by means of time series methods. In recent years, time series methods have found an increasing interest in medical statistics, in particular with regard to modelling and forecasting epidemiological series;⁵⁻⁸ in addition, they have been found useful in assessing relations between series which are dominated by seasonal variation and trends.^{9,10} Recently, the 'change-point problem' has become of epidemiological relevance when studying changing prevalence rates of hypopspadias (a rare malformation) by fitting a change-point Poisson process to registry data.¹¹ To our knowledge, time series methods have not been used to assess changes in health after an environmental disaster. In the present case study the problem of 'change' is studied using complementary time series methods. In the next section, a comparison between a forecast arising from models identified in the preaccident period (24 November 1985-31 October 1986) and the actual data is performed.¹² Thereafter, parsimonious intervention models¹³ which adequately represent the characteristic properties of the change are presented. Since some readers may be less familiar with intervention models this section is preceded by a short non-technical introduction. Finally, identification of a possible change-point is performed. In the following, we concentrate on the series of SYMPTOMS since only here were significant changes identified after the environmental accident.

STATISTICAL METHODS AND RESULTS

We let $\dots, y_{t-1}, y_t, y_{t+1}, \dots$ be the observations (entries, number of symptoms) at times $\dots, t-1, t, t+1, \dots$. Then it is assumed that y_t may be represented by an autoregressive integrated moving average model (ARIMA model).¹⁴

$$w_t = \nabla^d y_t,$$

and

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q},$$

$$\nabla y_t = y_t - y_{t-1}, \quad \nabla^2 y_t = \nabla(\nabla y_t).$$

The a_t are independent identically distributed random variables with expectation 0 and variance σ_a^2 . ϕ_1, \dots, ϕ_p are the autoregressive parameters, $\theta_1, \dots, \theta_q$ are the moving average parameters; ∇ is the differencing operator, d is the order of the differencing operator (usually $d = 0, 1$ or 2).

A short non-technical introduction to ARIMA models may be found in Reference 15. Extensive mathematical presentations of time series methods sufficient for the present case study are given in References 16 and 17. A condensed presentation of the application of time series methods in epidemiology may be found in Reference 18.

Comparison of forecast and actual data

A natural way to assess a possible change of a time series after an environmental accident is to compare actual data with a forecast arising from stochastic models which are appropriate before the accident.¹²

For the preaccident period of the series of SYMPTOMS, the iterative method of model identification, fitting and diagnostic checking proposed by Box and Jenkins¹⁴ was straightforward. The autocorrelation function showed an approximately exponential decay and the partial autocorrelation function showed a marked peak at lag 1. Therefore, an AR(1) model was

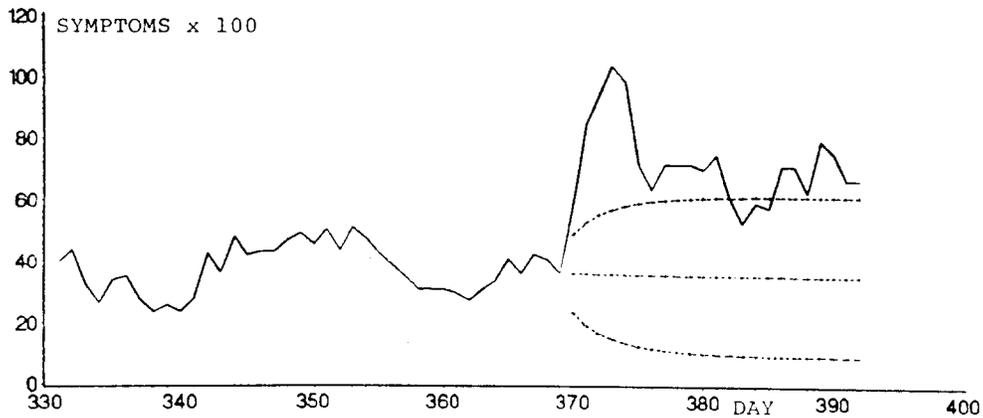


Figure 2. Solid curve: Last part of the series of SYMPTOMS ($\times 100$). Dashed curves: Estimated forecast function with 95 per cent probability limits

tentatively fitted. The autocorrelation function of the residuals showed no marked peaks. The Ljung-Box goodness-of-fit test¹⁹ revealed no sign of model inadequacy.

Figure 2 depicts the estimated forecast function arising from the preintervention model. According to the AR(1) model, it corresponds to an exponential curve starting at the last observed value (31 October 86) and finally reaching the mean value of the preaccident series. Since this last value is close to the mean value, the forecast function is hardly distinguishable from a straight line. In addition, the figure shows the 95 per cent probability limits for the individual forecasts, spanning a region where the actual data are expected to be found. The series of SYMPTOMS clearly exceeds the expected level.

An overall test by Tiao, Box and Hamming²⁰ may be applied to investigate the deviation between forecast and actual data. For lead time L , a statistic Q which follows approximately a χ^2 distribution with L degrees of freedom is calculated, where

$$Q = \sigma_a^{-2} \sum_{i=1}^L a_i^2$$

and the a_i are obtained by solving

$$e_l = \sum_{j=1}^l \psi_{l-j} a_j \quad (l = 1, \dots, L)$$

where $e_l = y_{T+l} - \hat{y}_T(l)$ are the forecast errors, T is the time of the accident, $\hat{y}_T(l)$ are the lead l forecasts at origin T and ψ_k are the weights of the moving average representation of the process:

$$y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

Here a value $Q = 83.2$ was found, which exceeds the 0.0001 value of χ^2 with $L = 23$ degrees of freedom.

More insight into the characteristic properties of the change may be found by means of intervention models presented in the next section.

Intervention models

Since some readers may be less familiar with intervention models¹³ we precede the discussion of results concerning 'Schweizerhalle' with a short non-technical introduction to this method.

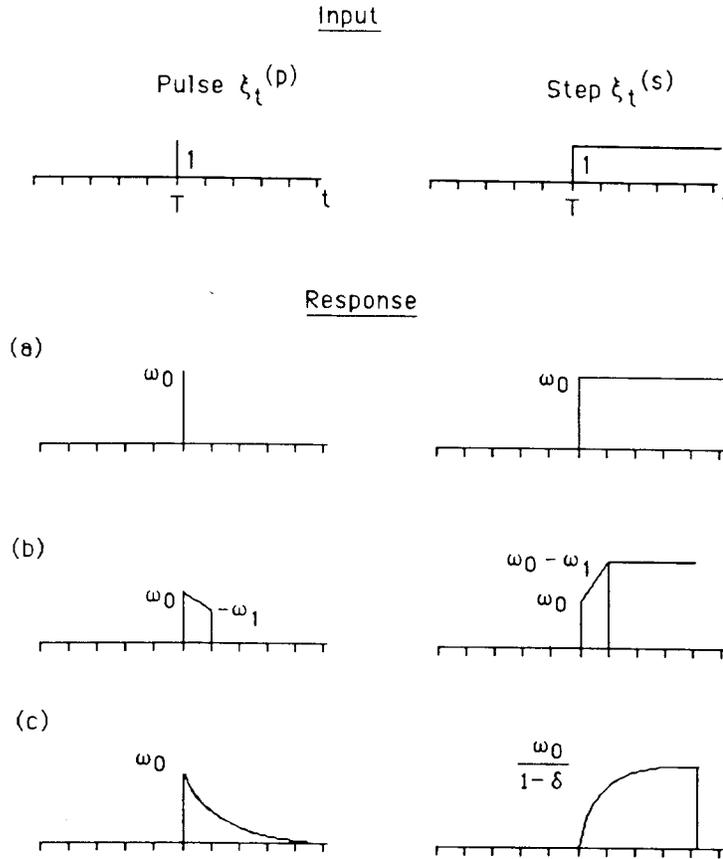


Figure 3. Basic patterns of response to a unit pulse and step input. For explanation see the section on Intervention models

The intervention model has the following structure:

$$y_t = u_t + n_t$$

where u_t represents the 'explained' part of the model and n_t represents the 'noise'. n_t is an ARIMA process as described above. The explained part u_t is considered to be the 'response' of a system to a dummy input variable ξ_t :

$$u_t = f(\xi_t).$$

The 'input' is usually taken as the unit step function $\xi_t^{(s)}$ or the unit pulse function $\xi_t^{(p)}$:

$$\xi_t^{(s)} = \begin{cases} 1 & \text{for } t \geq T \\ 0 & \text{for } t < T, \end{cases} \quad \text{and} \quad \xi_t^{(p)} = \begin{cases} 1 & \text{for } t = T \\ 0 & \text{otherwise.} \end{cases}$$

The step function $\xi_t^{(s)}$ may represent a disturbance starting at time T (the time of the accident). The pulse function $\xi_t^{(p)}$ may represent an event which acts only at time T .

Figure 3 shows basic types of response to the dummy input variables $\xi_t^{(s)}$ and $\xi_t^{(p)}$. Figure 3(a)

shows the simplest case where the response is just the input multiplied by a constant ω_0 :

$$f(\xi_t) = \omega_0 \xi_t.$$

This model predicts that the final level is reached immediately. A more refined model is shown in Figure 3(b):

$$f(\xi_t) = \omega_0 \xi_t - \omega_1 \xi_{t-1}$$

or

$$f(\xi_t) = (\omega_0 - \omega_1 B) \xi_t$$

where the symbol B is the so-called 'backshift' operator: $B \xi_t = \xi_{t-1}$. In this model the final level is reached in two steps. If $\omega_0 = 0$ the response is as in (a) but 1 time unit delayed.

Figure 3(c) shows another basic type of response where the final level is reached only gradually in the following way:

$$\begin{aligned} f(\xi_t) &= \omega_0(\xi_t + \delta \xi_{t-1} + \delta^2 \xi_{t-2} + \dots) \\ &= \omega_0(\xi_t + \delta B \xi_t + \delta^2 B^2 \xi_t + \dots) \\ &= \omega_0(1 + \delta B + \delta^2 B^2 + \dots) \xi_t \\ &= [\omega_0 / (1 - \delta B)] \xi_t. \end{aligned}$$

The above responses are special cases of the general response

$$f(\xi_t) = [\omega(B) / \delta(B)] \xi_t,$$

with

$$\begin{aligned} \omega(B) &= \omega_0 - \omega_1 B - \dots - \omega_s B^s, \\ \delta(B) &= 1 - \delta_1 B - \dots - \delta_r B^r. \end{aligned}$$

and the corresponding intervention model (of order r) is given by:

$$y_t = f(\xi_t) + n_t.$$

In the following, three intervention models of increasing complexity are fitted to the series of SYMPTOMS in such a way that each additional parameter allows a refined explanation of the data. The simplest model is found by setting $r = s = 0$:

- (i) intervention model of order 0:

$$y_t = \omega_0 \xi_t + n_t,$$

where ξ_t is the unit step function. The estimated parameters of this model are shown in Table I below the univariate model. Figure 4(a) shows as the second line the estimated function $\omega_0 \xi_t$ (the point labelled 370 corresponds to the date of the accident). This model gives a better fit to the data ($\sigma_a^2 = 0.00454$) than the univariate model. However, this simplified model does not fully represent all characteristic properties of the series: it predicts, for example, that the final level is reached immediately. It is therefore natural to consider the more elaborate model of order 1.

- (ii) intervention model of order 1:

$$y_t = \omega_0(1 - \delta_1 B)^{-1} \xi_t + n_t.$$

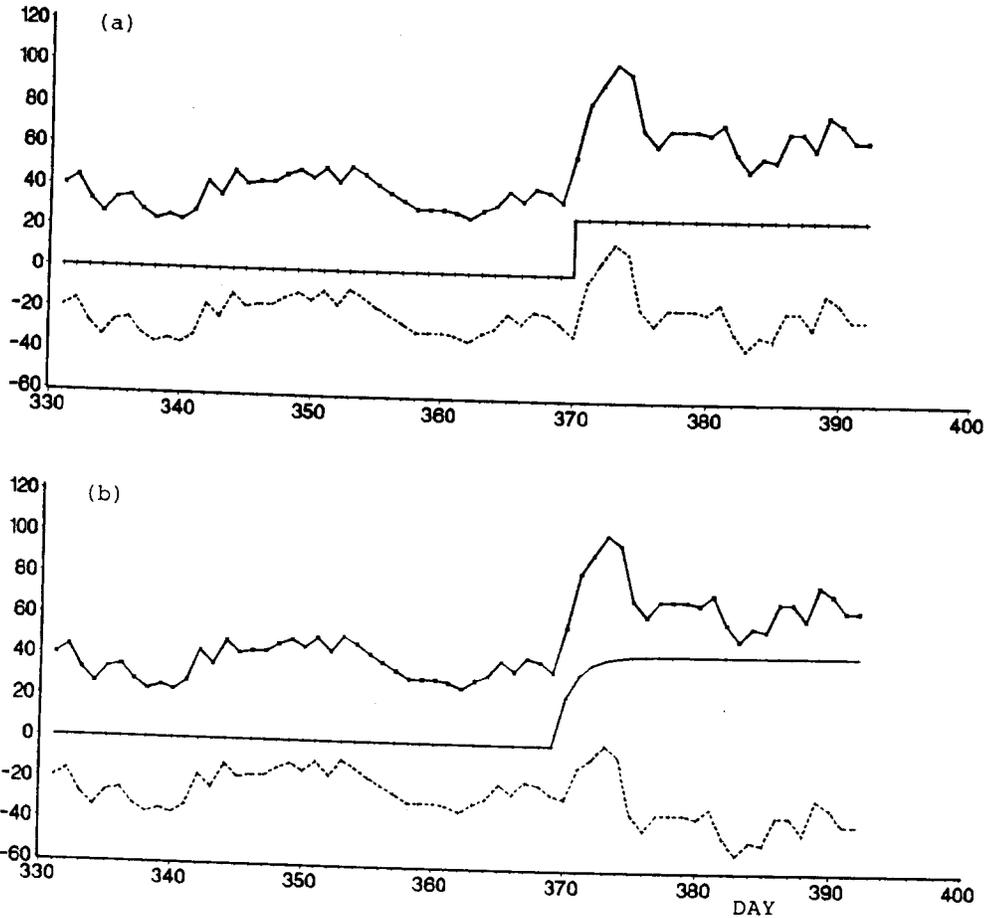


Figure 4. Two preliminary intervention models: (a) Model of order 0. (b) Model of order 1. Upper curve: Series of SYMPTOMS y_t . Second curve: $u_t = (\omega(B)/\delta(B))\xi_t$. Lower part: Noise series $n_t = y_t - u_t$ (shifted downwards). All curves are shown multiplied by 100

The parameters are given in Table I and the corresponding curves are presented in Figure 4(b). The residual variance decreases to $\sigma_a^2 = 0.00441$. This model allows for a gradual increase to the final level (middle curve of Figure 4(b)). However, the additional parameter δ_1 has a relatively large standard error. In addition, we recognize from the lowest curve of Figure 4(b) that the noise series n_t still has an unexplained 'bump'. This suggests the introduction of an additional refinement with a model of order 2.

(iii) intervention model of order 2:

$$y_t = \omega_0(1 - \delta_1 B - \delta_2 B^2)^{-1} \xi_t + n_t.$$

The 'explained' part of the model $u_t = \omega_0(1 - \delta_1 B - \delta_2 B^2)^{-1} \xi_t$ may be rewritten:

$$(1 - \delta_1 B - \delta_2 B^2)u_t = \omega_0 \xi_t$$

or

$$u_t - \delta_1 u_{t-1} - \delta_2 u_{t-2} = \omega_0 \xi_t$$

Table I. Summary of intervention models for the series of SYMPTOMS

Model type	Estimated parameters (SE)	Residual variance
Univariate	$\phi_1 = 0.91 (0.02)$ $\mu = 0.368 (0.039)$	0.00476
Intervention (order 0)	$\phi_1 = 0.87 (0.03)$ $\mu = 0.343 (0.026)$ $\omega_0 = 0.274 (0.058)$	0.00454
Intervention (order 1)	$\phi_1 = 0.87 (0.03)$ $\mu = 0.331 (0.027)$ $\omega_0 = 0.239 (0.065)$ $\delta_1 = 0.46 (0.16)$	0.00441
Intervention (order 2)	$\phi_1 = 0.87 (0.03)$ $\mu = 0.334 (0.026)$ $\omega_0 = 0.203 (0.033)$ $\delta_1 = 1.21 (0.07)$ $\delta_2 = -0.75 (0.07)$	0.00420

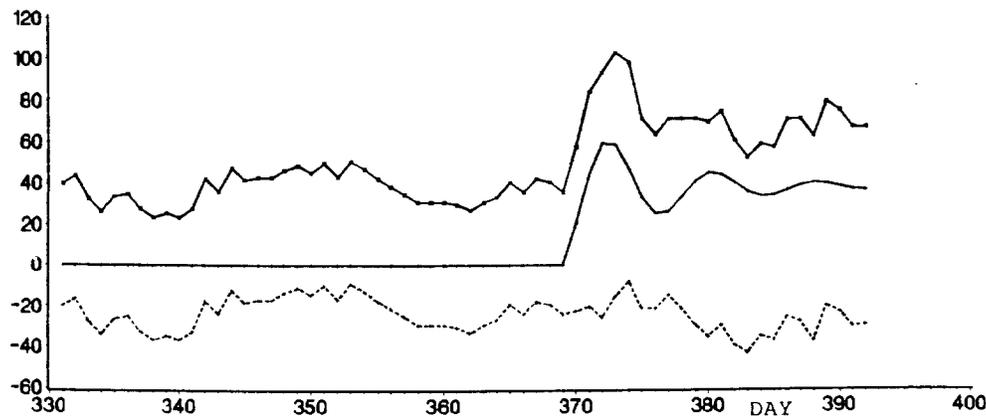


Figure 5. The final model: Intervention model of order 2. The same arrangement as in Figure 4

or

$$u_t = \delta_1 u_{t-1} + \delta_2 u_{t-2} + \omega_0 \xi_t.$$

This second-order difference equation may represent vibrations of discrete systems (in analogy to the differential equations of order 2 in continuous physical systems). The lowest part of Table I shows the estimated parameters of this model. The parameters ω_0 , δ_1 and δ_2 have small standard errors and the residual variance drops to $\sigma_a^2 = 0.00420$. The 'bump' in the noise series n_t (lowest curve of Figure 4) has disappeared, see Figure 5.

The intermediate curve u_t shows the characteristic behaviour of a 'damped vibration'. The final increase of the series over the level of the preaccident period is estimated by the gain $g = \omega_0(1 - \delta_1 - \delta_2)^{-1} = 0.376$, that is, an increase of approximately 0.38 respiratory symptoms per child per day.

Table II. Parts of the sequences of $MSE(T)$, $\sigma_a^2(T)$ and $p(T)$ which lie in the neighbourhood of the actual time of the accident (1 November 1986)

Date T	28.10	29.10	30.10	31.10	1.11	2.11	3.11	4.11	5.11
$MSE(T) (\times 10^4)$	183.2	180.5	178.8	176.6	173.2	174.3	181.0	189.6	199.8
$\sigma_a^2(T) (\times 10^4)$	49.3	49.1	47.9	46.8	44.8	46.2	48.2	47.9	48.9
$p(T)$	4×10^{-5}	5×10^{-4}	3×10^{-3}	3×10^{-2}	1	3.3×10^{-1}	3×10^{-4}	8×10^{-8}	6×10^{-12}

The introduction of additional parameters into the model ('overfitting'²¹) does not reduce the residual variance any further. The change of the series may thus be represented adequately and parsimoniously by an intervention model of order 2.

While the above results were found for the series of SYMPTOMS, no significant change was detected for the series of ENTRIES.

Identification of a change-point

In the present study, the time of the environmental accident is well known. However, we may assume that this time is unknown and ask whether the data themselves can tell us something about such a possible time. In the present context this might reveal a delay of possible effect. The identification of a 'change-point' for an ARIMA-process may be very difficult, both theoretically and with regard to numerical computation.²² However, it is possible to perform such identification, either by using exploratory approaches or by making simplifying assumptions about the underlying stochastic process.

A simple exploratory method to find this 'unknown' time consists in calculating a sequence of 2-sample t -tests (analysis of variance models), one for each possible time. The most likely time of the accident is the time T which gives the best model, that is, the model with the smallest error mean square (T was included in the second part of the series). The second line of Table II shows the part of the sequence of error mean squares $MSE(T)$ which lies in the neighbourhood of the actual time of the accident (1 November 1986). The absolute minimum of the sequence is found with T at 1 November 1986, which coincides with the actual date of the accident.

A more elaborate exploratory method assumes that there is better knowledge about the response of the system to a step input and about the correlation structure of the series. It makes use of the Box-Tiao intervention analysis. Instead of keeping the intervention time fixed, the step intervention function is scanned through all time periods. This method has been used before to identify the time at which a preventive measure may show an effect.²³

In the present context, we consider the following sequence of models (of order 2):

$$y_t = \omega_0^{(T)}(1 - \delta_1^{(T)}B - \delta_2^{(T)}B^2)^{-1} \xi_t^{(T)} + (1 - B)^{-1} a_t^{(T)},$$

$$\xi_t^{(T)} = \begin{cases} 1 & \text{for } t \geq T \\ 0 & \text{for } t < T, \end{cases}$$

where T runs through all possible times.

The third line of Table II shows the sequence of residual variances $\sigma_a^2(T)$. 'Box-Tiao scanning' identifies the same date as the simpler 't-test scanning'. However, the value of $MSE(T)$ at 1 November 1986 is quite close to the value of $MSE(T)$ at 2 November 1986. By contrast, the negative peak of $\sigma_a^2(T)$ is clearly isolated.

Finally, line 4 of Table II shows results obtained from a Bayesian approach: $p(T)$ is proportional to the posterior mass function of the change-point. It was calculated under the simplifying

assumption of a simple shift in mean in a normal sequence of independent random variables with vague prior information about the parameters. The precise formulation of the assumptions and the equations necessary for the calculation of the posterior mass function of the change point are given in Broemeling (pp. 308–312).²² This method identifies the same change-point as above. However, $p(T)$ decreases faster when moving away from 1 November 1986 than $MSE(T)$ or $\sigma_a^2(T)$. It is interesting to observe that $MSE(T)$ and $p(T)$ change less one step to the right than one step to the left of the change-point.

DISCUSSION

Possible health effects from environmental accidents constitute a basic threat to populations. It is therefore desirable to examine medical, or other available data, before and after an accident, from different points of view to gain a better understanding of a possible threat. A natural way to assess a possible change in a time series after an environmental accident is to compare actual data with a forecast arising from stochastic models which are appropriate before the accident. For the series of SYMPTOMS, an overall test showed a significant deviation of the actual post-accident series from the forecast ($P < 0.0001$).

A more quantitative understanding of the nature of the change is gained by means of intervention models. The identification of a sequence of models of increasing complexity showed that the response of the series of SYMPTOMS after the accident at 'Schweizerhalle' may be represented parsimoniously by an intervention model of second order. The model corresponds to what is known in continuous physical systems as a 'damped vibration'; after an initial overshoot, the series settles down to a new equilibrium at a higher level.

Common to these statistical methods is precise knowledge of the time of the accident. However, we may also assume that this time is unknown and ask if the data themselves tell us something about such a possible time. In studying the possible effect of preventive programmes it has been observed that the expected 'effect' may precede the actual time of 'intervention'.²³ It is possible as well that the effect shows up only after a certain delay.

Two exploratory methods, 't-test-scanning' and 'Box-Tiao scanning' were used to identify a possible change-point. For the series of SYMPTOMS, both methods found 1 November 1986 to be the most likely date, which coincides with the actual time of the accident.

Finally a possible change-point was identified using a Bayesian approach,²² with the same results as above. However, it was found that $MSE(T)$ and $p(T)$ change less one step to the right than one step to the left of the change-point. This is probably due to the fact that the change is only incompletely characterized by a simple immediate increase in mean. As discussed above it may be represented adequately by an intervention model of second order.

All results obtained strongly support an increase in the number of symptoms per child after the accident. Yet, these results should be interpreted with caution. Even though no direct hints of epidemics exist (from serological investigations performed by paediatricians), the influence of viral infections cannot be entirely excluded. The identified intervention model states that after an initial overshoot following the accident the series settles down to a new level (0.38 units higher than the preaccident mean). Unfortunately, data were not available for a longer period after the accident; thus, there is a possibility that 'return to normal' could have been missed. The question of whether the accident caused more symptoms to be recorded than were actually present cannot be answered entirely satisfactorily. However, other studies conducted in this context point toward an increase in respiratory symptoms in the general population.²⁴

Another problem concerns the possible presence of seasonal effects. The well known ARIMA models which allow modelling of random seasonal effects, require data over several years for their

identification. If a series containing marked seasonal variation is observed only over one year, pronounced signs of non-stationarity would be expected. In particular, when fitting an AR(1) model the corresponding autoregressive coefficient, ϕ_1 , should be approximately equal to one, indicating the need for differencing. In the present case it was found that $\phi_1 = 0.91$ (SE: 0.02) (Table I, first line). In addition, the residual variance of the competing non-stationary random walk model (ARIMA(0, 1, 0) model) was $\sigma_a^2 = 0.00496$, slightly larger than that of the AR(1) model (Table I). Thus, it is interesting to note that the '1-year' series of SYMPTOMS may be considered a stationary series with strong autocorrelation (and fixed mean) rather than a non-stationary series.

In the present study, the series of ENTRIES did not show an increase after the accident. A possible explanation might be that symptoms experienced were too mild to warrant consultation. In addition, the accident occurred on a Saturday and media informed the population that any symptoms they might experience would soon disappear. This might have prevented people from visiting a paediatrician or an outpatients' clinic and partly explain the apparent difference between the two time series. The finding suggests an examination of population based data, as well as medical records, in estimating health effects of environmental accidents.

We recommend that readers interested in political aspects of the accident look at the short note in the Lancet.²⁵

ACKNOWLEDGEMENT

The authors thank the referees for their encouraging and constructive help.

REFERENCES

1. Braun-Fahrlander, C., Ackermann-Lieblich, U., Wanner, H. U., Rutishauser, M., Gnehm, H. P. and Minder, C. H. 'Auswirkungen von Luftschadstoffen auf die Atemwege von Kleinkindern', *Schweizerische Medizinische Wochenschrift*, **119**, 1424-1433 (1989).
2. Helfenstein, U., Ackermann-Lieblich, U., Braun-Fahrlander, C. and Wanner, H. U. 'Air pollution and diseases of the respiratory tracts in pre-school children: a transfer function model', *Journal of Environmental Monitoring and Assessment*, **17**, (1991) in press.
3. Anon. 'Europe's river Rhine has suffered one of the worst chemical spills ever', *Environmental Science & Technology*, **21**, 5 (1987).
4. Capel, P.D., Giger, W. and Wanner, O. 'Accidental input of pesticides into the Rhine river', *Environmental Science & Technology*, **22**, 992-997 (1988).
5. Katzoff, M. 'The application of time series forecasting methods to an estimation problem using provisional mortality statistics', *Statistics in Medicine*, **8**, 335-341 (1989).
6. Schnell, D., Zaidi, A. and Reynolds, G., 'A time series analysis of gonorrhoea surveillance data', *Statistics in Medicine*, **8**, 343-352 (1989).
7. Zaidi, A., Schnell, D. and Reynolds, G. 'Time series analysis of syphilis surveillance data', *Statistics in Medicine*, **8**, 353-362 (1989).
8. Martinez-Schnell, B. and Zaidi, A. 'Time series analysis of injuries', *Statistics in Medicine*, **8**, 1497-1508 (1989).
9. Bowie, C. and Prothero, D. 'Finding causes of seasonal diseases using time series analysis', *International Journal of Epidemiology*, **10**, 87-92 (1981).
10. Helfenstein, U. 'Detecting hidden relations between time series of mortality rates', *Methods of Information in Medicine*, **29**, 57-60 (1990).
11. Simpkin, J. M. and Downham, D. Y. 'Testing for a change-point in registry data with an example on hypospadias', *Statistics in Medicine*, **7**, 387-393 (1988).
12. Box, G. E. P. and Tiao, G. C. 'Comparison of forecast and actuality', *Applied Statistics*, **25**, 195-200 (1976).
13. Box, G. E. P. and Tiao, G. C. 'Intervention analysis with applications to economic and environmental problems', *Journal of the American Statistical Association*, **70**, 70-79 (1975).

14. Box, George E. P. and Jenkins, Gwilym M. *Time Series Analysis, Forecasting and Control*, revised edition, Holden Day, San Francisco, 1976.
15. Helfenstein, U. 'Box-Jenkins modelling of some viral infectious diseases', *Statistics in Medicine*, **5**, 37-47 (1986).
16. Abraham, B. and Ledolter, J. *Statistical Methods for Forecasting*, Wiley, New York, 1983.
17. Chatfield, C. *The Analysis of Time Series: An Introduction*, Chapman and Hall, London, 1989.
18. Helfenstein, U. 'The use of transfer function models, intervention analysis and related time series methods in epidemiology', *International Journal of Epidemiology*, **20** (1991), in press.
19. Ljung, G. M. and Box, G. E. P. 'On a measure of lack of fit in time series models', *Biometrika*, **65**, 297-303 (1978).
20. Tiao, G. C., Box, G. E. P. and Hamming, W. J. 'Analysis of Los Angeles photochemical smog data: a statistical overview', *Journal of the Air Pollution Control Association*, **25**, 260-268 (1975).
21. Jenkins, Gwilym M. *Practical Experiences with Modelling and Forecasting Time Series*, Gwilym Jenkins & Partners (Overseas) Ltd., St. Helier, 1979.
22. Broemeling, L. D. *Bayesian Analysis of Linear Models*, Marcel Dekker, New York, 1985.
23. Helfenstein, U. 'When did a reduced speed limit show an effect? Exploratory identification of an intervention time', *Accident Analysis & Prevention*, **22**, 79-87 (1990).
24. Ackermann-Liebrich, U., Braun-Fahrlander, C. and Rapp-Pestalozzi, R. 'Epidemiologic analysis of an environmental disaster - the 'Schweizerhalle' experience', manuscript submitted.
25. Anon. 'A fishy smell in Basle', *Lancet*, **1**, 95 (1987).