Stochastic Environmental Research and Risk Assessment 16 (2002) 112–126 © Springer-Verlag 2002
DOI 10.1007/s00477-002-0090-8

Accuracy of spatio-temporal RARX model predictions of water table depths

M. Knotters, M. F. P. Bierkens

Abstract. Time series of water table depths (H_t) are predicted in space using a regionalised autoregressive exogenous variable (RARX) model with precipitation surplus (P_t) as input variable. Because of their physical basis, RARX model parameters can be guessed from auxiliary information such as a digital elevation model (DEM), digital topographic maps and digitally stored soil profile descriptions. Three different approaches to regionalising RARX parameters are used. In the 'direct' method (DM) P_t is transformed into H_t using the guessed RARX parameters. In the 'indirect' method (IM) the predictions from DM are corrected for observed systematic errors. In the Kalman filter approach the parameters of regionalisation functions for the RARX model parameters are optimised using observations on H_t. These regionalisation functions describe the dependence on spatial co-ordinates of the RARX parameters. External drift kriging and simple kriging with varying means are applied as regionalisation functions, using guessed RARX model parameters or DEM data as secondary variables. Predictions of H_t at given days, as well as estimates of expected water table depths are made for a study area of 1375 ha. The performance of the three approaches is tested by cross-validation using observed values of H_t in 27 wells which are positioned following a stratified random sampling design. IM performs significantly better with respect to systematic errors than the alternative methods in estimating expected water table depths. The Kalman filter methods perform better than both DM and IM in predicting the temporal variation of H_t, as is indicated by lower random errors. Particularly the Kalman filter method that uses DEM data as an external drift outperforms the alternative methods with respect to the prediction of the temporal variation of the water table depth.

Keywords: DEM, groundwater, Kalman filtering, spatio-temporal uncertainty, validation.

1 Introduction

Agricultural and ecological water management in the Netherlands requires accurate information on water table fluctuation. Preferably, the uncertainty about

M. Knotters (✉), M. F. P. Bierkens
Alterra, PO Box 47, 6700 AA Wageningen, The Netherlands
e-mail: m.knotters@alterra.wag-ur.nl

This study was funded by Research Program 328 of the Netherlandish Ministry of Agriculture, Nature Management and Fisheries. We are grateful to an anonymous referee for his helpful comments.

water table fluctuation is quantified, in order to enable risk assessment. The information may concern either prediction of the actual water table depth on a given day, or estimation of statistics. These statistics describe the dynamics of the water table. They are for instance expected water table depths on a given day in any future year, or the probability that the water table depth is within a critical depth, for instance at the start of the growing season (Knotters and Bierkens, 2001). Predictions of actual water table depths on a given day are more and more required in the current water management practice in the Netherlands. By using predictions of actual water table depths, water managers can anticipate actual trends. For instance, in the wet spring of 1998 up-to-date predictions of water table depths, for instance at March 13, 1998, might have been helpful to water managers in aiming for optimal conditions at the start of the growing season. In contrast, statistics of the water table fluctuation are needed to assess options for long term water policy. For instance, to decide on investments to lower the water table depth in the start of the growing season, an estimate of the expected water table depth at March 13 in any future year, given the prevailing hydrological regime, may be relevant information. Agricultural or ecological water management in the Netherlands is generally restricted to areas of limited size, in which generally only a few time series of water table depth are available. On average, 1 suitable observation well per 750–1250 ha is present (Finke, 2000). Therefore, it may be attractive to use auxiliary information in predicting the water table depths in these areas. Knotters and Bierkens (2001) introduced a regionalised autoregressive exogenous variable model (RARX) for the relationship between precipitation surplus and water table depth. In essence, the RARX model is a linear time series model which parameters are made space dependent or 'regionalised'. In other words, the value of a RARX parameter depends on its location. For locations at which time series of water table depths were observed, the RARX model parameters can be calibrated. At other locations the RARX model parameters can be guessed from auxiliary physical information which is generally widely available, as explained in detail by Knotters and Bierkens (2001). This auxiliary information includes a digital elevation map (DEM), digitally stored soil profile descriptions, soil physical standard curves and digital information on the positions and sizes of ditches.

Using the guessed RARX parameter fields it is possible to predict time series of water table depth spatially. This can be done straightforwardly, without use of observed water table depths. However, applying this 'direct' method can result in large systematic prediction errors (Knotters and Bierkens, 2001). The direct method can be improved by interpolating the observed systematic prediction errors to unvisited locations in order to correct the preliminary results. Furthermore, if repeated observations in time are available, the noise term, containing the part of water table fluctuation that cannot be explained from the precipitation surplus, can be quantified. Knotters and Bierkens (2001) showed that this 'indirect' method can be applied in estimating the risks of shallow water tables at the start of the growing season.

Bierkens et al. (2001) incorporated the RARX model into a space-time Kalman filter algorithm. This application of the space-time Kalman filter is a mathematical framework which enables to predict water table depths in space and time conditionally to observed water table depths, optionally using auxiliary information. For every time step at which the water table depth has been observed an update of the predictions is made. This makes the Kalman filter attractive in practising daily water management. As the 'indirect' method, the Kalman filter approach can also be used for estimation of statistics of the water table fluctuation.

The aim of this study is to evaluate the accuracy of prediction methods in either predicting actual water table depths, or estimating expected water table depths in any future year in the prevailing hydrological regime. To this end a cross-validation experiment is carried out. The accuracy of estimates of expected water table depths ‘in any future year’ is evaluated as follows: expected water table depths are estimated for days in the monitoring period at which observations were taken, given the meteorological conditions during the monitoring period. Next, the estimates are compared with the observations.

The setup of the paper is as follows. In Sect. 2 the study area and the data set are introduced. In Sect. 3 the ‘direct’ and ‘indirect’ spatiotemporal prediction method, and three applications of the space-time Kalman filter algorithm are explained. An outline of the validation procedure is given in Sect. 4. The results are presented and discussed in Sect. 5. The paper ends with some conclusions.

2
Study area; data set

The study area of 1375 ha is situated in the north-eastern part of the Netherlands. The set of observed water table depths H_t (Fig. 1) can be divided into two parts:

1. a preferential sample of 13 time series observed in permanently installed observation wells (measurement locations);
2. a stratified random sample (approximately proportionally allocated to the stratum areas) of 27 short time series (2 years length) observed in temporarily installed observation wells (validation locations), representing a part of the study area of 976 ha. The four strata are catchments with controlled surface water levels, see Fig. 1.

The northern ‘appendix’ of the study area (Fig. 1) is not a part of the area of 976 ha from which stratified the random sample has been taken. In the northern

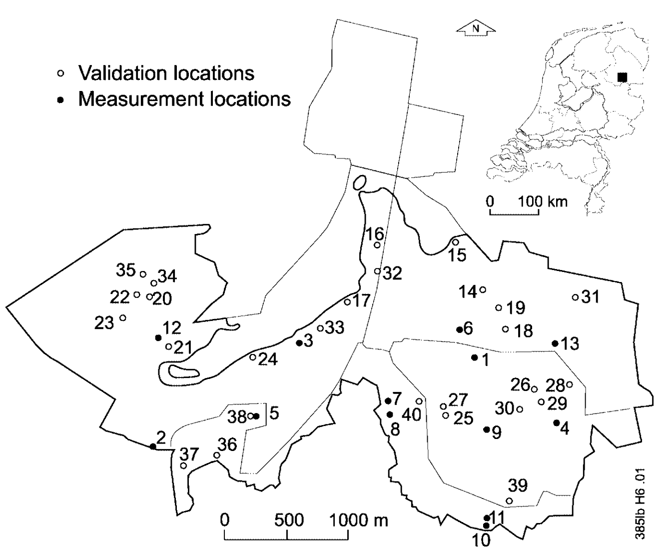

Fig. 1. Study area, its location in the Netherlands and its division into catchments. Closed circles: 13 locations where time series of water table depth have been observed in permanently installed wells. Open circles: 27 locations where water table depths have been observed incidentally in temporarily installed observation wells

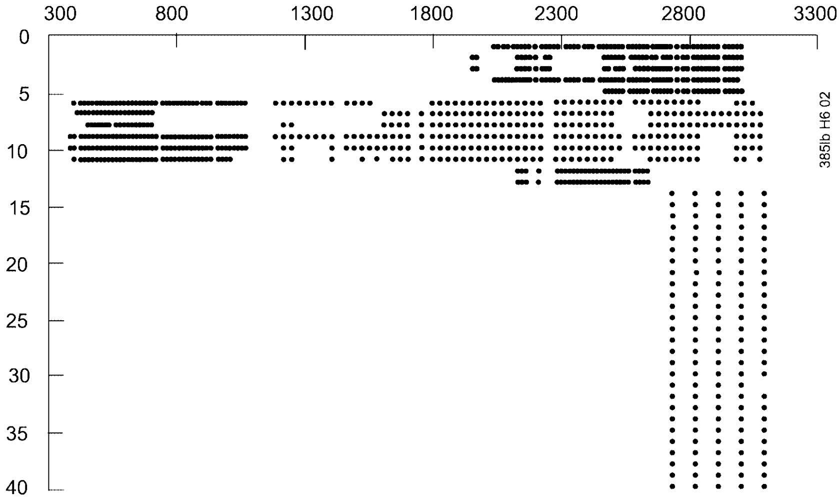

Fig. 2. Sampling scheme. Vertical axis: well number. Horizontal axis: day number (day 1 = January 1st, 1990)

part boulder clay is present at shallow depth. Because of the occurrence of temporal water tables at impermeable layers, the water table depth can not be measured accurately in observation wells placed in the boulder clay. Therefore, the northern part of the study area was excluded in the validation. Predicted water table depths are mapped for the northern part, however, in order to illustrate the differences in the predictions obtained by the evaluated methods.

The ground surface elevation was determined at all 40 groundwater observation points. Figure 2 shows the sampling scheme in time. Note that the interval lengths are not constant.

The soil of the study area consists of relatively wet and peaty sediments in a brook valley in the southern part, and relatively dry cover sands, locally covering boulder clays in the northern part. A digital elevation model (DEM, Fig. 3) reflects the geomorphologic structure of the landscape: a relatively flat and low brook valley in the south and relatively high, slightly undulating cover sand ridges in the north.

Daily data on precipitation are available from the station Dedemsvaart nearby the study area. Daily data on the potential Makkink crop reference evapotranspiration (Winter et al., 1995) are available from the station Eelde at approx. 55 km distance of the study area. The precipitation surplus is calculated as follows:

$$P_t = P_t^* - E_{p,t} \ ,$$

where P_t^* is the daily precipitation $[\mathrm{L}t^{-1}]$ (measured at 9.00 h AM) and $E_{p,t}$ is the daily potential Makkink crop reference evapotranspiration $[\mathrm{L}t^{-1}]$ (measured at 0.00 h). Note that $P_t^*, E_{p,t}$ and $H_t(\mathbf{u})$ were not measured at the same time. It has been experienced in practice, however, that this will not lead to serious errors. It is important to note that the water table depths are generally observed after 9.00 h AM, that is, after the observation of daily precipitation.

At 1185 locations in the study area soil profile descriptions were made by experienced soil surveyors, as a part of a regional survey for land development. A

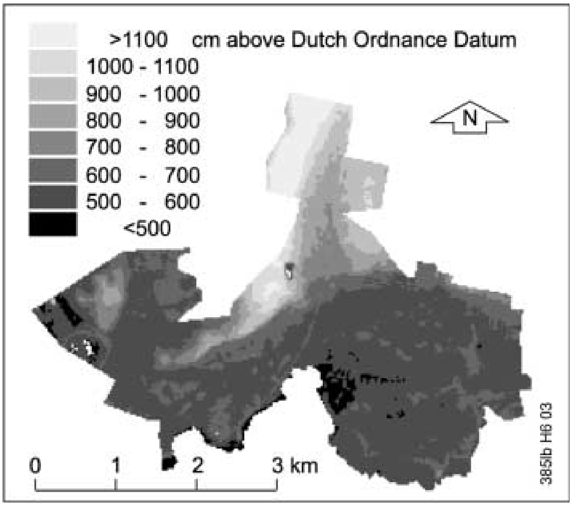

Fig. 3. Digital elevation model of the study area. Values in cm above Dutch Ordnance Datum

DEM, made by laser scanning, is available with a resolution of 25 × 25 m, see Fig. 3. On a point scale, the DEM deviates on average 22 cm from the elevation determined by surveying. Information on ditches, such as average distances and perimeters, are derived from digital topographic maps. Information on surface water levels is provided by the local water authorities.

3
Spatio-temporal prediction methods

3.1
The RARX model

Knotters and Bierkens (2001) explain the RARX model for water table depths in detail. Here the model will be recapitulated. The RARX($\mathbf{u}$;1,0) model is given by

$$H_t(\mathbf{u}) - \mu(\mathbf{u}) = a_1(\mathbf{u})\{H_{t-\Delta t}(\mathbf{u}) - \mu(\mathbf{u})\} + b_0(\mathbf{u})P_t + \epsilon_t(\mathbf{u}) \quad , \tag{1}$$

with $\mathbf{u}^{\mathrm{T}} = (x, y)$ indicating the spatial co-ordinates x and y. Δt is the time step. Note that the precipitation surplus P_t is assumed to be global, that is, space invariant. This assumption is reasonable for relatively small areas. The error process, $\epsilon_t(\mathbf{u})$, is assumed to form a discrete sequence of mutually independent disturbances with zero expectation and finite and constant variance $\sigma^2_\epsilon(\mathbf{u})$. In space, the error process is assumed to be continuous and structured.

According to Knotters and Bierkens (2000) the physical basis of the RARX($\mathbf{u}$;1,0) model parameters is derived from the water balance of a soil column. The RARX($\mathbf{u}$;1,0) model parameters can be expressed in physical quantities as follows:

$$a_1(\mathbf{u}) = \mathrm{e}^{-\Delta t/(\varphi(\mathbf{u})\gamma(\mathbf{u}))} \quad , \tag{2}$$

$$b_0(\mathbf{u}) = \gamma(\mathbf{u})\{1 - a_1(\mathbf{u})\} \quad , \tag{3}$$

and

$$\mu(\mathbf{u}) = \gamma(\mathbf{u})q_{\mathrm{b}}(\mathbf{u}) + H_{\mathrm{s}}(\mathbf{u}) \quad , \tag{4}$$

$\varphi(\mathbf{u})$ is the effective porosity of the soil in which the water table fluctuates [-]. $\gamma(\mathbf{u})$ is the drainage resistance [t], which is defined as

$$\gamma(\mathbf{u}) = \frac{H_t(\mathbf{u}) - H_{\mathrm{s}}(\mathbf{u})}{q_{\mathrm{d}}(\mathbf{u})} \quad ,$$

$H_{\mathrm{s}}(\mathbf{u})$ is the drainage level (e.g. surface water level) [L], and $q_{\mathrm{d},t}(\mathbf{u})$ is the drainage flux $[\mathrm{Lt}^{-1}]$. $q_{\mathrm{b}}(\mathbf{u})$ is the flux to the shallow groundwater from the deeper groundwater systems $[\mathrm{Lt}^{-1}]$. $\varphi(\mathbf{u}), \gamma(\mathbf{u}), q_{\mathrm{b}}(\mathbf{u})$, and $H_{\mathrm{s}}(\mathbf{u})$ are assumed to be time invariant. Furthermore, it is assumed that the water table fluctuation depends on precipitation surplus only. Other influences such as groundwater discharge are assumed to be absent or forming part of the time invariant flux $q_{\mathrm{b}}(\mathbf{u})$.

The physical relationship for $\sigma^2_\epsilon(\mathbf{u})$ is:

$$\sigma^2_\epsilon(\mathbf{u}) = \mathrm{var}\left\{ b_0(\mathbf{u})\left[(E_{\mathrm{p},t} - E_{\mathrm{a},t}(\mathbf{u})) - \frac{\Delta V(\mathbf{u})}{\Delta t}\right] + \epsilon^*_t(\mathbf{u}) \right\} \quad , \tag{5}$$

where $E_{\mathrm{a},t}(\mathbf{u}))$ is the actual evapotranspiration $[\mathrm{Lt}^{-1}]$ and $V(\mathrm{u})$ is the moisture volume in the unsaturated zone. Both quantities can not be observed against reasonable costs and therefore generally are unknown. The term $\epsilon^*_t(\mathbf{u})$ contains the remaining unknown influences as well as the model error. Thus, to estimate $\sigma^2_\epsilon(\mathbf{u})$ measured water table depths are needed.

The RARX($\mathbf{u}$;1,0) model in (1) is the basis of the three prediction methods that are evaluated in this study: the 'direct' method, the 'indirect' method and the space-time Kalman filter.

3.2
The 'direct' method (DM) and the 'indirect' method (IM)

In DM widely available auxiliary information is transformed directly into RARX parameters. Knotters and Bierkens (2001) describe how $\gamma(\mathbf{u})$ and $\varphi(\mathbf{u})$ are guessed from auxiliary information on ground surface elevation, drainage devices and soil profile descriptions. Using the relationships in (2) and (3), the guessed physical parameters $\tilde{\gamma}(\mathbf{u})$ and $\tilde{\varphi}(\mathbf{u})$ are transformed in guessed RARX parameters $\tilde{a}_1(\mathbf{u})$ and $\tilde{b}_0(\mathbf{u})$. The parameter $\mu(\mathbf{u})$ is guessed directly from hydromorphic soil characteristics which were determined at the 1185 augering locations. Equation (4) was not used in guessing $\mu(\mathbf{u})$, because it was not possible to guess the regional groundwater flux $q_{\mathrm{b}}(\mathbf{u})$ accurately from auxiliary information.

The RARX parameters can be guessed for locations where soil profile descriptions were made. Next, the guessed parameter values are interpolated to unvisited locations (e.g. a grid). Using these interpolated, physically based RARX parameters $\tilde{a}_1(\mathbf{u}), \tilde{b}_0(\mathbf{u}), \tilde{\mu}(\mathbf{u})$ and the model given in (1), time series on precipitation surplus are transformed into time series on water table depths. Note that DM is a purely deterministic method; observed water table depths are not used. As shown by Knotters and Bierkens (2001) DM may result in large systematic prediction errors. Furthermore, a prediction of the standard deviation of the error process, $\sigma_\epsilon(\mathbf{u})$ is not provided, which makes the direct method unsuitable for application in risk assessment.

As in DM, IM starts with transformation of physical information into RARX parameters and subsequently, after spatial interpolation, transformation of time series on precipitation surplus into time series on water table depths. Next, the predicted water table depths are compared with observed water table depths. The observed systematic errors are spatially interpolated to correct the preliminary, direct, predictions. In this study, the systematic errors were interpolated by external drift kriging (Deutsch and Journel, 1992; Goovaerts, 1997), using elevation data from the DEM as an external drift.

If time series of water table depths are observed, the standard deviation of the error process, $\sigma_\epsilon(\mathbf{u})$, can be estimated. The estimated $\sigma_\epsilon(\mathbf{u})$ can be interpolated spatially to unvisited locations, and thus it can be used in drawing maps reflecting the risk that a critical water table depth is exceeded (Knotters and Bierkens, 2001).

3.3
The space–time Kalman filter (KF)

In DM, observed water table depths are not used, whereas in IM observed water table depths are only used to correct afterwards for systematic errors, and to predict the variance of the error term, $\sigma_\epsilon(\mathbf{u})$. However, spatio-temporal predictions of water table depths can possibly be improved when they are optimised with respect to observed water table depths. The space–time Kalman filter algorithm enables one to predict water table depths in space and time, conditional to the observed water table depths. Furthermore, the space–time Kalman filter can be used in conditional simulation and in network optimisation. Bierkens et al. (2001) described in detail the Kalman filter algorithm used in this study. The basis of the algorithm is the RARX($\mathbf{u}$;1,0) model given in (1). The parameters $a_1(\mathbf{u}), b_0(\mathbf{u}), \mu(\mathbf{u})$, and $\sigma_\epsilon(\mathbf{u})$ can be calibrated for those locations where time series of water table depth are available which are sufficiently long to calibrate an ARX(1,0) model. At all other locations, the parameters have to be estimated by so called regionalisation functions. These functions describe the dependence on spatial co-ordinates of the RARX model parameters. Since the number of time series is generally small in study areas and thus the number of known (calibrated) RARX parameters is limited, these regionalisation functions should be able to include auxiliary information. An estimator of a RARX($\mathbf{u}$;1,0) parameter, for instance $\hat{a}_1$, is given by

$$\hat{a}_1(\mathbf{u}) = r_a[\mathbf{u}; a_1(\mathbf{u}_i), i = 1, \ldots, n; s_j(\mathbf{u}), j = 1, \ldots, m; \boldsymbol{\theta}_a] \ , \qquad (6)$$

(Bierkens et al., 2001), where $a_1(\mathbf{u}_i)$ are parameter values calibrated on time series observed at n locations $\mathbf{u}_i, i = 1, \ldots, n$, and $s_j(\mathbf{u}), j = 1, \ldots, m$ are the values of m auxiliary variables at location $\mathbf{u}$. In this study, $m = 1$, i.e., only one auxiliary variable is considered. $\boldsymbol{\theta}_a$ are the parameters of the regionalisation function. For instance, if some type of kriging is used as a regionalisation function, $\boldsymbol{\theta}_a$ are the variogram parameters.

In the example of the RARX parameter $a_1(\mathbf{u}), s_j(\mathbf{u})$ can for instance be a guessed value of $a_1(\mathbf{u})$, derived from physical auxiliary information using (2). If an interpolation method such as kriging is used for $r_a(\cdot)$, then calibrated values of $a_1(\mathbf{u})$ are required and thus sufficiently long time series of water table depths must be available. In this case study 13 suitable time series are available.

Two types of regionalisation function that account for an auxiliary variable were applied: kriging with an external drift, and simple kriging with varying means (Goovaerts, 1997; Deutsch and Journel, 1992). In kriging with an external drift the regionalisation function for the RARX parameter $a_1(\mathbf{u})$ has the following form:

$$r_a(\mathbf{u};\ldots) = \beta[s(\mathbf{u})] + \sum_{i=1}^{n} \lambda_i(\boldsymbol{\theta}_a)\{a_1(\mathbf{u}_i) - \beta[s(\mathbf{u}_i)]\} \ , \tag{7}$$

where $\beta(\cdot)$ is a drift function depending on the auxiliary variable $s(\mathbf{u}), \lambda(\boldsymbol{\theta}_a)$ are kriging weights depending on the parameter vector $\boldsymbol{\theta}_a$. Here $\boldsymbol{\theta}_a$ is the vector of parameters of the variogram of the residuals. The drift function has a linear form:

$$\beta[s(\mathbf{u})] = b'_0 + b'_1 s(\mathbf{u}) \tag{8}$$

The RARX parameters $\tilde{a}_1(\mathbf{u}), \tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$ which were guessed on the basis of physical information (Sect. 3.2) are used as an external drift. As explained in Sect. 3.1, $\sigma_\epsilon(\mathbf{u})$ can not easily be guessed from physical auxiliary information against reasonable costs. However, it can be expected that a relationship exists between $\sigma_\epsilon(\mathbf{u})$ and the average water table depth. This is pointed out as follows. The effective porosity will increase with increasing water table depth, because the fraction of pores filled with air will increase. Consequently, the value of the term $b_0(\mathbf{u})$ in (5) will decrease, as follows from (2) and (3). Thus, it can be expected that $\sigma_\epsilon(\mathbf{u})$ is related to the average water table depth. Therefore, the values of $\tilde{\mu}(\mathbf{u})$ which were guessed from hydromorphic soil characteristics roughly reflecting the average water table depth, can be useful as auxiliary information in regionalising $\sigma_\epsilon(\mathbf{u})$. The parameter $\sigma_\epsilon(\mathbf{u})$ was regionalised using external drift kriging with $\tilde{\mu}(\mathbf{u})$ as an external drift.

In simple kriging with varying means the regionalisation function is in its structure equal to the function in (7). However, the methods have different definitions of the trend component, $\beta[\cdot]$. In external drift kriging, the trend component is a simple linear regression function, see (8). The parameters b'_0 and b'_1 in (8) are implicitly estimated through the kriging system within each search neighbourhood. In contrast, in simple kriging with varying means the trend can have any form. the trend, or 'varying mean', can for instance be the result of arbitrary ad-hoc procedures, whereas in external drift kriging inference relies on the generalised least squares method. In this study, the 'varying mean' is formed by guessed RARX parameter values, which can be considered as the outcomes of ad-hoc procedures. Thus, for the RARX parameter $a_1(\mathbf{u})$, the trend component $\beta[\tilde{a}_1(\mathbf{u})]$ is defined as

$$\beta[\tilde{a}_1(\mathbf{u})] = \tilde{a}_1(\mathbf{u})$$

in simple kriging with varying means, whereas in external drift kriging the trend component is defined as

$$\beta[\tilde{a}_1(\mathbf{u})] = b'_0 + b'_1\tilde{a}_1(\mathbf{u}) \ ,$$

where b'_0 and b'_1 are estimated through the kriging system.

3.4
Summary of prediction methods

The method described in Sects. 3.2 and 3.3 require different minimum sizes of ground-water data sets. DM does not need observed water table depths at all. IM needs observed water table depths at a sufficiently large number of points in space, in order to accurately estimate the model of spatial structure of the systematic prediction error. Preferably, time series on water table depths are observed at these points so that the standard deviation of the error process $\sigma_\epsilon(\mathbf{u})$ can be estimated and next spatially predicted at unvisited locations. If in KF a kriging interpolation method is used as a regionalisation function, calibrated ARX(1,0) parameters are needed at the well locations. This implies that sufficiently long time series on water table depth must be available. Additional incidental observations are needed to calibrate the regionalisation parameters $\boldsymbol{\theta}_a$ (Bierkens et al., 2001).

Both in DM and IM auxiliary information on drainage resistance, effective porosity and hydromorphic soil characteristics is needed, as is explained in Sect. 3.2. In KF the RARX model parameters that are guessed following the same procedure as in DM can be used as auxiliary variables. Alternatively, in situations where the physical information needed for guessing RARX model parameters is absent, other auxiliary information such as a DEM can be used in KF. The prediction methods evaluated in this study are now summarized as follows:

- DM: the direct method. Physically based, guessed RARX parameters are used in the predictions, observed water table depths are not used;
- IM: the indirect method. Physically based, guessed RARX parameters and all observed water table depths are used;
- KF1: space–time Kalman filter, with external drift kriging as a regionalisation function for $a_1(\mathbf{u}), b_0(\mathbf{u}), \mu(\mathbf{u})$. Physically based, guessed RARX parameters are used as external drifts. $\sigma_\epsilon(\mathbf{u})$ is regionalised by external drift kriging with $\tilde{\mu}(\mathbf{u})$ as an external drift. All observed water table depths are used;
- KF2: space–time Kalman filter, with simple kriging with varying means as a regionalisation function for $a_1(\mathbf{u}), b_0(\mathbf{u}), \mu(\mathbf{u})$. Physically based, guessed RARX parameters are used as varying means. $\sigma_\epsilon(\mathbf{u})$ is regionalised by external drift kriging with $\tilde{\mu}(\mathbf{u})$ as an external drift. All observed water table depths are used;
- KF2: space–time Kalman filter, with external drift kriging as a regionalisation function for $a_1(\mathbf{u}), b_0(\mathbf{u}), \mu(\mathbf{u})$, and $\sigma_\epsilon(\mathbf{u})$. The ground surface elevation from the DEM is used as auxiliary variable. All observed water table depths are used.

Two types of information are distinguished: predicted water table depths, and estimates of statistics for the water table fluctuation, for instance expected water table depths. In DM and IM, one the same procedure is followed for both predicting water table depths and estimating expected water table depths: in both cases the input series is transformed in an output series, using guessed RARX parameter values. In IM, observed water table depths are only used to correct for systematic errors, both in prediction and in estimation. However, if KF methods are used for predicting water table depths, for each time step at which observations on water table depths are available an update of the spatiotemporal predictions is made. Alternatively, if KF is used for estimation of expected water table depths, first the RARX($\mathbf{u}$;1,0) parameters are estimated using the available observations on water table depths and the auxiliary information used in the regionalisation function. Next, time series on precipitation surplus are transformed into time series on water table depth by using the estimated RARX($\mathbf{u}$;1,0) parameters, without updating.

4
Set up of the validation

Water table depths were predicted (or estimated) at the 27 validation locations by means of a cross-validation procedure. In IM and the KF methods all observed water table depths were used in the prediction (see Figs. 1 and 2), except those observed at the selected validation location. This procedure was repeated 27 times, each time leaving one validation location out for which prediction errors were calculated. The interval length in (1) is one day ($\Delta t = 1$ day) in the prediction procedures, since daily precipitation surplus is the input of the RARX model.

For each test location the following validation measures were calculated:

- the mean error,

$$ME_{hi} = \frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} e_{j,hi} \quad , \tag{9}$$

where hi indicates the ith test location in the hth stratum, e is the difference between observed and predicted water table depth resulting from the cross-validation, n_{hi} is the length of the observed time series at location hi. The absolute value, $|ME|_{hi}$ is a measure for the closeness of the predicted to the observed mean water table depth;

- the standard deviation of error,

$$SDE_{hi} = \sqrt{\frac{1}{n_{hi} - 1} \sum_{j=1}^{n_{hi}} [e_{j,hi} - ME_{hi}]^2} \quad , \tag{10}$$

which is a measure for the closeness of predicted to observed temporal fluctuation;

- and the root mean squared error,

$$RMSE_{hi} = \sqrt{\frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} e^2_{j,hi}} \quad , \tag{11}$$

which is a measure for the overall closeness of predicted to observed water table depths.

The three measures in (9)–(11) describe the closeness of predicted time series of water table depth to observed time series at location $\mathbf{u}_{hi}$. Areal means of these measures can be estimated by

$$m_A(y) = \sum_{h=1}^{L} W_h \left[\frac{\sum_{i=1}^{n_h} y_{hi}}{n_h} \right] \quad , \tag{12}$$

(Cochran, 1977), where L denotes the number of strata, and W_h denotes the stratum weight. The stratum weight is proportion of the stratum area to the total area. y can be replaced by ME, $|ME|$, SDE or $RMSE$.

5 Results and discussion

5.1 Predicting water table depths

The areal means of ME, $|ME|$, SDE, and $RMSE$ in predicting water table depths are listed in Table 1. The SDE can be interpreted as the closeness of the predicted to the observed temporal fluctuation. Table 1 shows that KF methods perform well

Table 1. Validation results for the prediction of water table depths

	DM	IM	KF1	KF2	KF3
$m_A(ME)$	−23.6	4.5	8.3	−10.0	14.4
$m_A(\|ME\|)$	30.7	17.5	24.1	28.1	24.2
$m_A(SDE)$	9.5	9.5	6.9	7.0	5.7
$m_A(RMSE)$	33.5	21.5	26.1	29.6	25.9

with respect to *SDE*. In particular KF3 performs well with respect to *SDE*, despite the use of only DEM data as auxiliary variable, whereas in the alternative methods all available physical information was used.

The systematic errors (*ME*) are relatively large for all methods except IM. As explained in Sect. 3.2, correction for systematic errors is part of the prediction procedure in IM. The negative *ME* values found for DM and KF2 can to some extent be explained from possible bias in the guessed RARX parameter $\tilde{\mu}(\mathbf{u})$. The parameter $\tilde{\mu}(\mathbf{u})$ has been guessed from hydromorphic soil characteristics observed in the field by augering at 1185 locations as part of a soil survey. $\tilde{\mu}(\mathbf{u})$ is calculated by the average of the depth to the top of the permanently reduced zone and the depth to the top of the hydromorphic characteristics, such as rust mottles. This average value should approximate the mean water table depth. However, hydromorphic characteristics are often 'fossil', that is, the hydromorphic characteristics do not represent the actual water table, but reflect the water table in the very past which generally fluctuated at shallower depths. Besides this, the mean water table depth differs from the RARX parameter $\mu(\mathbf{u})$ by a factor $b_0(\mathbf{u})\bar{P}/(1 - a_1(\mathbf{u}))$, where $\bar{P}$ is the mean precipitation surplus. The contribution of this factor to the negative systematic errors is generally limited, however. Bias in $\tilde{\mu}(\mathbf{u})$ will be reflected in the predictions of KF2, because here $\tilde{\mu}(\mathbf{u})$ is used as an auxiliary variable in simple kriging with varying means (see Sect. 3.3). In KF1, however, $\tilde{\mu}(\mathbf{u})$ is used as an auxiliary variable in kriging with an external drift. In this method the values of the auxiliary variable are linearly transformed by a regression model (Eq. 8), which possibly eliminates bias in $\tilde{\mu}(\mathbf{u})$.

Figure 4 shows maps of the predicted actual water table depth at March 13, 1998, as an example of predicting actual water table depths. The relatively detailed patterns resulting from IM and KF3 can be explained by the DEM data from a 25× 25 m grid which were used as an external drift variable in both methods. It is interesting to compare the patterns obtained by DM with the patterns obtained by IM, KF1, and KF2, because differences are the effect of using observed water table depths in the predictions. Most resemblance is found for DM and KF2, particularly in the northern part of the study area. In Sect. 2 it is explained that in this northern 'appendix' the water table depth was not observed. The relatively small differences between DM and KF2 indicate that the observed water table depths contribute less to the predictions in KF2 than in IM and KFI. Clearly, the stochastic components of the regionalised RARX parameters, for example $\{a_1(\mathbf{u}_i) - \beta[s(\mathbf{u}_i)]\}$ in (7), are small in KF2. In KF1 kriging with an external drift is used as a regionalisation function. Here, the drift $\beta[s(\mathbf{u}_i)]$ is a linear transformation of the guessed RARX parameter values, see (7) and (8). The patterns in Fig. 4 indicate that in KF1, with kriging with an external drift as a regionalisation function, the observed water table depths contribute more to the predictions than if kriging with varying means is used as a regionalisation function in KF2.

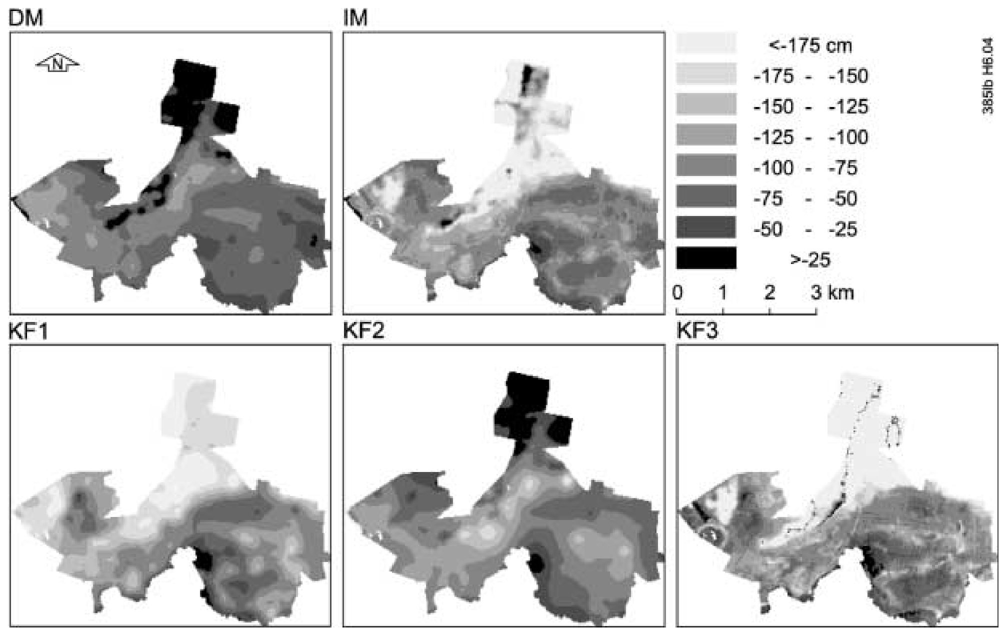

Fig. 4. Predicted water table depth at March 13, 1998. Values in cm below ground surface. DM: direct method, IM: indirect method, KF1: space–time Kalman Filter with guessed RARX parameters as an external drift, KF2: space–time Kalman Filter with guessed RARX parameters as varying means, KF3: space–time Kalman Filter with DEM data as an external drift

Table 2. Validation results for the estimation of expected water table depths

	DM	IM	KF1	KF2	KF3
$m_A(ME)$	−23.6	4.5	16.3	−8.8	25.0
$m_A(\|ME\|)$	30.7	17.5	30.2	29.1	32.3
$m_A(SDE)$	9.5	9.5	9.5	9.3	9.9
$m_A(RMSE)$	33.5	21.5	32.6	31.8	35.0

5.2
Estimating expected water table depths

Figure 5 shows maps of the expected water table depth at March 13 in any future year, given the prevailing hydrological regime. The future meteorological conditions were approximated by the daily precipitation surplus data from 1969 to 1998. As compared to the actual water table depths at March 13, 1998 (Fig. 4), the water table at March 13 in any future year (Fig. 5) is predicted at larger depths. Indeed, this result reflects the relatively wet meteorological conditions in the Netherlands in the springs of 1998. Again, resemblance is found for the maps obtained by DM and KF2.

In Table 2 the areal means of *ME*, |*ME*|, *SDE*, and *RMSE* for estimates of expected water table depths are given. Again negative *ME* values are found for DM and KF2; the possible causes are discussed in Sect. 5.1. As compared to prediction of actual water table depths (Table 1), the performance of KF methods decreases in estimating expected water table depths by using KF methods. This can be explained from the fact that in this case no updates are made for each time step at which observations become available. Instead, the time series on precipitation surplus is transformed straightforwardly into a time series of estimates of expected water table depths, by using the preliminary calibrated RARX(**u**;1,0) parameters.

The results in Table 2 indicate that the methods perform equally well with respect to *SDE*. The results on |*ME*| in Table 2 indicate that IM approximates the mean water table depth more than KF1 and KF3. In particular KF3 results in large systematic errors. Note that in KF3 only DEM data used as auxiliary variable, whereas in the alternative methods all available physical information was incorporated.

6
Conclusions

The prediction performance of five methods based on the RARX(**u**;1,0)model for water table depths was tested in a case study. The performance was evaluated by means of a cross-validation procedure, since the number of observed time series of water table depths was too small to separate an independent validation set. A disadvantage of cross-validation is that the distances from the observation points to the prediction points are overrated, because an incomplete set of observation points is used in the interpolation. Thus the prediction errors may be overestimated. For this reason the cross-validation results are mainly used to evaluate the relative prediction performance of the five methods.

Distinction was made between prediction of water table depths and estimation of expected water table depths. Predictions of water table depths are needed to support short term decisions in daily water management. It is concluded that the temporal fluctuation of the table is predicted more precisely by methods which

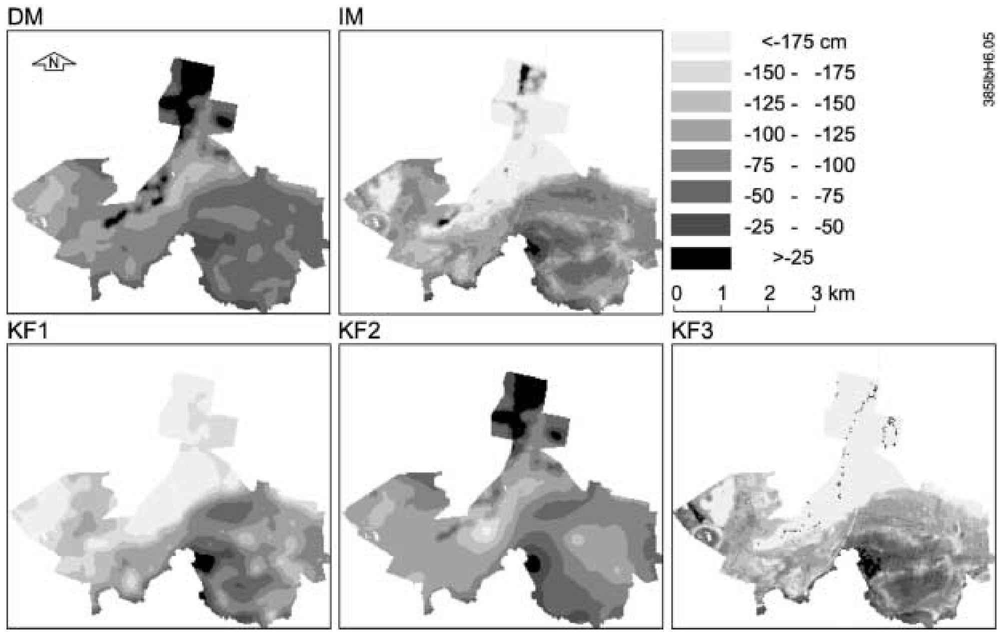

Fig. 5. Estimates of the water table depth, expected for March 13 in any future year, given the prevailing hydrological regime. DM: direct method, IM: indirect method, KF1: space–time Kalman Filter with guessed RARX parameters as an external drift, KF2: space–time Kalman Filter with guessed RARX parameters as varying means, KF3: space–time Kalman Filter with DEM data as an external drift

include the space–time Kalman filter algorithm than by the 'direct' and 'indirect' method which are based on guessed RARX parameter fields and geostatistically interpolated corrections for systematic error. This may be expected, because in the space–time Kalman filter methods an update of the predictions is made for every time step at which the water table depth has been observed. However, the Kalman filter methods fail in predicting the mean water table depth, whereas the indirect method predicts the mean level accurately.

Estimates of expected water table depths are needed to support decision making in long term water policy. The case study indicates that the indirect method (IM) is an accurate alternative for the Kalman filter methods in predicting water table depths which are expected in a given hydrological regime. In particular in the Kalman filter method that makes use of DEM data as an auxiliary variable in external drift kriging (KF3) relatively large systematic errors can occur.

Maps of predicted water table depths and estimates of expected water table depths indicate that observations contribute more to the predictions if kriging with an external drift is used as a regionalisation function for the RARX parameters in the space–time Kalman filter algorithm (KF1) than if simple kriging with varying means (KF2) is used.

References

Bierkens MFP, Knotters M, Hoogland T (2001) Space–time modelling of water table depth using a regionalized time series model and the Kalman filter. Water Resour. Res. 37: 1277–1290

Cochran WG (1977) Sampling Techniques. Wiley, New York

Deutsh CV, Journel AG (1992) GSLIB. Geostatistical Software Library and User's Guide, Oxford University Press, New York

Finke PA (2000) Updating the (1:50,000) Dutch groundwater table class map by statistical methods: an analysis of quality versus cost. Geoderma 97: 329–350

Goovaerts P (1997) Geostatistics for Natural Resources Evaluation. Oxford University Press, New York

Knotters M, Bierkens MFP (2000) Physical basis of time series models for water table depths. Water Resour. Res. 36: 181–188

Knotters M, Bierkens MFP (2001) Predicting water table depths in space and time using a regionalised time series model. Geoderma 103: 51–77

Winter TC, Rosenberry DO, Sturrock AM (1995) Evaluating of 11 equations for determining evapotranspiration for a small lake in the north central United States. Water Resour. Res. 31: 983–993