ARC Geophysical Research

Alternative Formulations of the Watertable Fluctuation Method of Recharge Estimation: A Quantitative Comparison

A. L. Becke[1,2], S. C. Solórzano-Rivas[1, 2], and A. D. Werner*[1, 2]

[1]College of Science and Engineering, Flinders University, Australia

[2]National Centre for Groundwater Research and Training, Flinders University, Australia

Abstract

Groundwater recharge estimation is vital for ensuring the sustainable management of water resources and the protection of ecosystems. The watertable fluctuation method (WTFM) is commonly used to estimate distributed groundwater recharge in unconfined aquifers. There are many variations in the method's implementation, mainly in the extrapolation of recession curves to estimate groundwater discharge during recharge events. Despite the popularity of the WTFM, the results of the most popular variations have not been compared to assess their accuracy. This study examines six alternative forms of the WTFM by applying them to 1,000 model-generated hydrographs to determine the most accurate approach. Recharge estimation error was characterised according to model input parameters: model input recharge, specific yield, aquifer length, transmissivity, distance between the observation well and the groundwater divide (relative to the aquifer length), and distance between the observation well and the groundwater discharge boundary. The RISE method, which does not account for ongoing discharge during recharge events, was the poorest performing, underestimating gross recharge by an average of 22% for the cases tested. The exponential local recession curve method (the most common approach) also tended to underestimate recharge (by an average of 14% for the cases tested). The fixed-timestep master recession curve method was the most accurate, underestimating recharge on average by 4% for the cases tested. This approach assumes a greater rate of discharge as the watertable rises, which is more consistent with Darcy's Law. The linear local recession curve was the second-most accurate method despite its simplicity, underestimating gross recharge by an average of 7% for the cases tested. The widest range of recharge estimation error occurred near the groundwater discharge boundary, for all variants tested. For those cases, a high transmissivity increased the likelihood of recharge underestimation. These findings provide valuable insights for improving the reliability of WTFM applications in groundwater recharge investigations.

Keywords: Groundwater; Aquifer; Hydrogeology; Master Recession Curve; Hydrograph Analysis

*Corresponding Author

1 Introduction

Groundwater recharge estimation is vital for ensuring the sustainable management of water resources, including the protection of ecosystems and the prediction of system reliability under increasing water demands. Recharge is difficult to directly measure given the spatial variability in soil and landscape properties and the transience of surface hydrological processes. This is compounded by the need to obtain recharge at scales relevant to water management areas. For this reason, Scanlon et al. [26] recommend applying multiple methods to estimate recharge because all approaches are accompanied by significant uncertainties. The most widely used approaches to estimate recharge are based on groundwater responses to recharge (and discharge), such as watertable fluctuations or the groundwater salinity, thereby providing recharge values that integrate complex infiltration and evapotranspiration processes over significant spatial and/or temporal scales [6] .

A widely used method to estimate distributed recharge in unconfined aquifers is the watertable fluctuation method (WTFM). The WTFM assumes that a watertable rise is due to infiltration from the land surface arriving at the watertable [15] rather than focused recharge derived from surface water bodies or lateral groundwater inflows from neighbouring aquifers. The WTFM can be expressed as:

$$R_{\mathrm{e}} = S_{\mathrm{y}} \frac{\Delta h_{\mathrm{r}}}{\Delta t} , \tag{1}$$

where R_{e} is the groundwater recharge [L T^{-1}] obtained from application of the WTFM; S_{y} is aquifer specific yield [-]; Δh_{r} is the change in watertable elevation attributable to recharge [L]; and Δt is the period over which recharge causes an observable change in the hydrograph [T], typically equal to the duration of watertable rise. The extraction of Δh_{r} from a groundwater hydrograph is shown in Figure 1.

A detailed discussion of the WTFM methodology and its limitations is provided by Becke et al. [3], who identified various methods for calculating Δh_{r} in Equation 1. Some applications of the WTFM consider only the groundwater-level rise (line CD, Figure 1) resulting from a recharge event in assigning values to Δh_{r} [e.g., 9]. This approach neglects the groundwater discharge occurring during periods of watertable rise due to, for example, evapotranspiration of groundwater, discharge to surface water bodies, lateral groundwater flow away from the recharge area, etc. When only the watertable rise is adopted for Δh_{r}, the 'net recharge' is obtained [14]. In most groundwater management applications, an estimate of gross recharge is required, for example, as an input into groundwater models [7]. The gross recharge is the sum of net recharge (line CD, Figure 1) and an approximation of the groundwater discharge (line DE, Figure 1) that occurred during the period of recharge. A common approach to estimate the latter is to project the antecedent recession curve forward to the time of the peak watertable elevation [3], as shown by line BE in Figure 1.

The projected recession curve is generally assumed to represent the decline in the watertable that would have occurred in the absence of recharge [9, 14, 18, 30, 34, 37]. The forward projection (line BE, Figure 1) of antecedent recession curves (line AB, Figure 1) for this purpose usually adopts exponential decay or power functions [e.g., 23, 25, 33]. This includes

E-mail address: adrian.werner@flinders.edu.au
doi: https://doi.org/10.5149/ARC-GR.1417

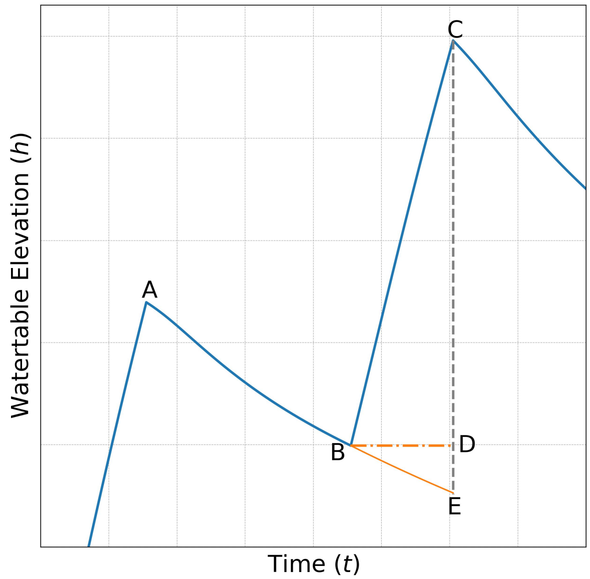


Figure 1: Schematic of a groundwater hydrograph (blue line) illustrating an application of the WTFM. The antecedent recession curve is represented by line AB, while line BE depicts the projection of line AB. The change in watertable elevation attributable to recharge (Δh_r) is delineated as line CE when considering gross recharge, or line CD for the estimation of the net recharge. Line DE represents a drop in the watertable elevation attributable to groundwater discharge. Line BD denotes the duration of watertable rise, presumed to coincide with the period of recharge (Δt).

WTFM applications that involve manual extrapolation (i.e., "graphical method") of the recession curve [3]. The rate of discharge obtained from these approaches, represented by the slope of line BE (Figure 1), decreases as the watertable rises. That is, the forward-extrapolation of the antecedent recession curve using an exponential decay function will tend towards a horizontal line (no change in groundwater discharge) with time, despite a rising watertable. This is in contradiction with general concepts of groundwater hydraulics, whereby discharge is proportional to the hydraulic gradient (i.e., in accordance with Darcy's Law), which is expected to increase as the watertable rises considering that discharge is usually associated with topographic low points of largely stable elevation in the landscape [e.g., 8, 29]. Linear extrapolation of the recession curve, such as that used by Wilopo and Putra [35], can be considered to represent a constant rate of discharge (during watertable rise), presuming that a watertable rise or fall is proportional to the associated storage change in the aquifer.

Fixed-timestep WTFM approaches estimate groundwater recharge (and discharge) as the difference between groundwater level changes and projected recession curves at fixed intervals [e.g., 16], as an alternative to the event-based approach [e.g., 27] shown in Figure 1. The method of application of fixed-timestep WTFM approaches leads to greater rates of discharge when the watertable is higher, even where exponential-decay curves are adopted in assessing groundwater discharge [5]. This is prima facie more consistent with Darcy's Law (i.e., higher groundwater heads lead to greater discharge to surface features of fixed elevation) than event-based WTFM approaches (i.e., Figure 1) that adopt exponential-decay curves. Whether or not a better estimate of recharge is obtained with these different time-stepping and recession extrapolation techniques has not been systematically assessed.

Prior attempts to compare alternative approaches of the WTFM against known recharge values include the study of Águila et al. [1], who conducted a parametric and numerical analysis of the uncertainties of the WTFM using two variants of the method: (a) the RISE method, which uses a fixed timestep and neglects the groundwater discharge occurring during periods of watertable rise (therefore estimating net recharge), and (b) the graphical method, which is event-based and involves manual projection of recession presuming an exponential-decay function. Águila et al. [1] applied the RISE and graphical methods to synthetic hydrographs created from numerical models. They found that the errors of the graphical approach were up to 15% less than those of the RISE method when compared to the known recharge input to the numerical model. The RISE method underestimated recharge (i.e., the gross recharge applied to the model) by up to 80% when the river water level remained constant during the recharge event. If the river stage rose simultaneously with the watertable during recharge events, and the monitoring location was close to the river (within 50 m), the RISE method over-estimated recharge by approximately 400% when stream-aquifer connectivity was high, and by 8% when stream-aquifer connectivity was low.

Solórzano-Rivas et al. [27] also applied the WTFM to synthetic hydrographs, which they created using an analytical solution for intermittent recharge. Two approaches were examined: an exponential Local Recession Curve (LRC) approach, and an exponential event-based Master Recession Curve (MRC) approach. LRC approaches consider only the recession behaviour immediately prior to individual groundwater-level rise events, while MRC approaches use a recession function that represents the average behaviour of multiple recession periods. As both approaches adopt exponential functions, the rate of discharge decreases as the watertable rises (for reasons explained above). Solórzano-Rivas et al. [27] examined the accuracy of these methods by applying them to 3,000 recharge events for which recharge was a-priori known. They concluded that the LRC and MRC approaches under-estimated recharge by an average of 13% and 14% respectively, highlighting the need to revise the approaches used to extrapolate the antecedent recession curve. Interestingly, the underestimation of recharge observed in the Solórzano-Rivas et al. [27] study contradicts the findings of numerous field studies, which conclude that the WTFM often overestimates recharge [e.g., 20, 32, 36]. Overestimation in field-based studies is often attributed to uncertainty in S_y [e.g., 5, 10] and the inclusion of watertable fluctuations unrelated to recharge in WTFM calculations [20, 32, 36]. However, there has been less focus on the uncertainties associated with ongoing drainage estimates.

Other studies have compared alternative forms of the WTFM in field applications [e.g., 9, 11, 21, 24]. For instance, Delin et al. [9] reported significant discrepancies, with deviations of up to 63%, between recharge estimates from a fixed-timestep MRC approach and the RISE method. Gumuła-Kawecka et al. [11] also compared a fixed-timestep MRC approach and the RISE method, finding that estimates derived from the former were more than twice those of the latter. Nimmo et al. [24] undertook a comparative analysis between a fixed-timestep MRC approach and an event-based MRC approach, revealing mismatch between the two that ranged from -37% (fixed-timestep MRC recharge < event-based MRC recharge) to 41%. Likewise, Lanini [21] observed that the RISE method consistently underestimated recharge by 12 to 50% compared to the event-based MRC approach, based on the application of these methods to a multitude of case studies.

The current study builds on these prior comparisons of WTFM approaches by quantitatively evaluating six commonly used variants of the WTFM to find the most accurate approach, thereby addressing the recommendations of Becke et al. [3] who concluded that the optimal approach for extrapolating the recession curve is unclear when implementing the WTFM. Solórzano-Rivas et al. [27] found that the two forms of the WTFM that they applied produced biased recharge estimates, although they tested only the event-based MRC and exponential LRC approaches, representing a small subset of the methods adopted in published cases. They concluded that approaches involving linear extrapolation of the recession curve, or functions with an increasingly negative gradient over time, merit further investigation. The WTFM variants compared in this study make differing assumptions regarding the rate of groundwater discharge that occurs during recharge events, and they have not previously been compared. WTFM variants are evaluated by applying them to synthetic hydrographs with known gross recharge and S_y, where the latter is constant in both time and depth. As such, we evaluate only the method of recession extrapolation, and omit an assessment of S_y, measurement error, and other factors contributing to uncertainty in WTFM applications. The practical challenges associated with implementing each WTFM variant were also assessed by applying the variants to data observed in the field. This paper extends the work of Águila et al. [1] and Solórzano-Rivas et al. [27] by comparing a wider array of WTFM approaches beyond the RISE and graphical methods of Águila et al. [1], and the exponential LRC and event-based MRC approaches of Solórzano-Rivas et al. [27].

2 Methods

2.1 Watertable Fluctuation Method Approaches

The six variants of the WTFM used in this study to estimate recharge are distinguished by the technique used to approximate the head change, Δh_r, that is adopted in Equation 1. These include the following approaches: (1) linear LRC, (2) power LRC, (3) exponential LRC, (4) event-based MRC, (5) fixed-timestep MRC, and (6) RISE method. Recharge was calculated using Equation 1 for all approaches. A summary of the approaches used is provided in Table 1, whilst each approach is depicted graphically in Figure 2.

Table 1: Summary of the WTFM approaches examined in the current study.

Category of recession analysis	**Name of method**	**Type of recharge estimated**	**Time-step**	**Groundwater discharge rate (with watertable rise)**	**Reference**
LRC	Linear LRC	Gross	Event-based	Constant	Wilopo and Putra [35]
	Power LRC	Gross	Event-based	Decreasing	Wendland et al. [34]
	Exponential LRC	Gross	Event-based	Decreasing	Solórzano-Rivas et al. [27]
MRC	Event-based MRC	Gross	Event-based	Decreasing	Nimmo et al. [24], Nimmo and Perkins [23]
	Fixed-timestep MRC	Gross	Fixed	Increasing	Heppner and Nimmo [16]
No recession projection	RISE	Net	Fixed	Not considered	Delin et al. [9]

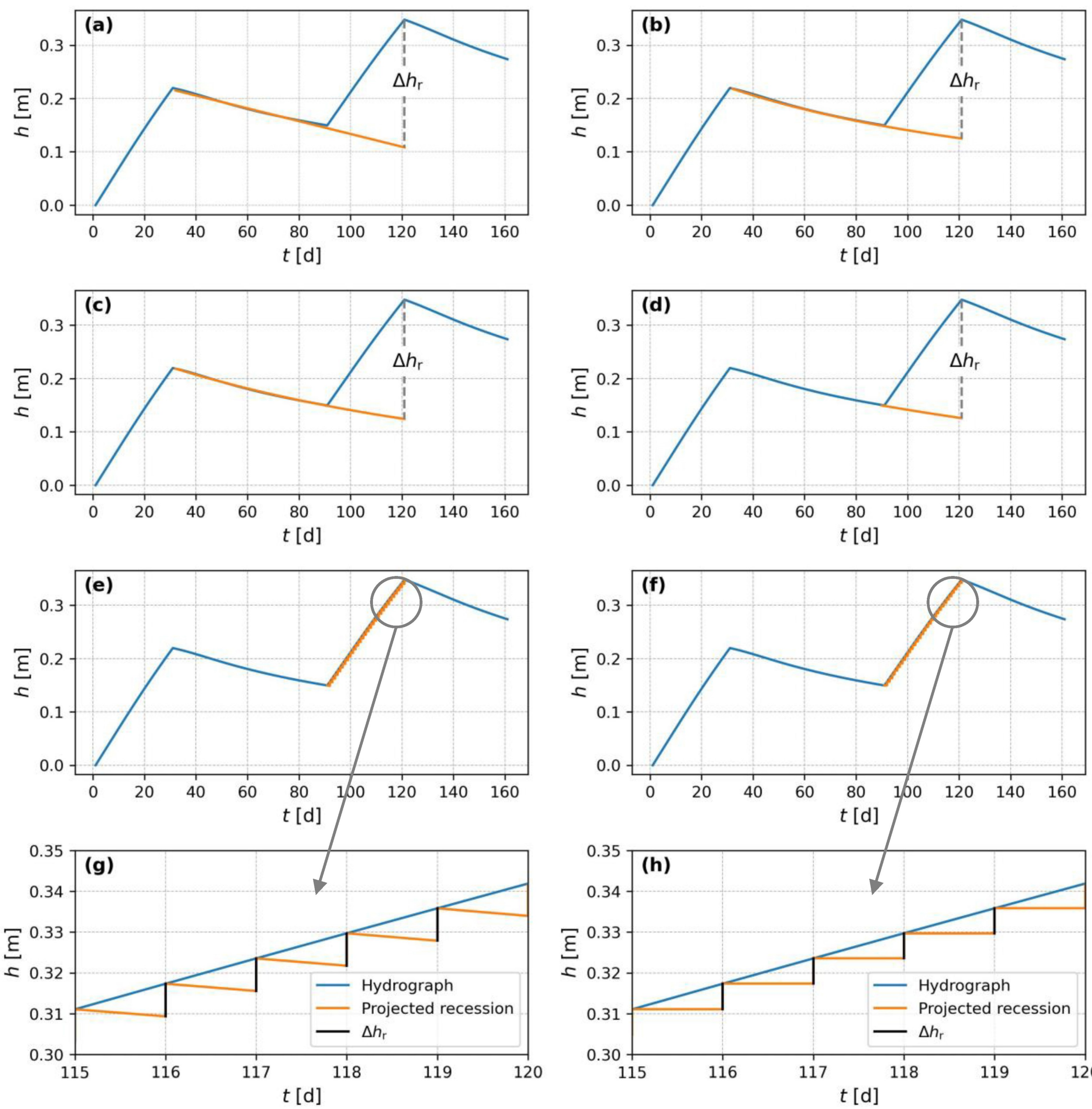


Figure 2: Graphical comparison of the WTFM approaches explored in this study: (a) linear LRC, (b) power LRC, (c) exponential LRC, (d) event-based MRC, (e) fixed-timestep MRC, (f) RISE approach, (g) close-up of (e), (h) close-up of (f).

Three variants of event-based LRC's were applied by fitting the antecedent recession curve (line AB, Figure 1) using alternative fitting functions to describe the rate of groundwater level change (dh/dt), which is necessary to construct the projected recession. For example, the linear LRC (Figure 2a), outlined by Wilopo and Putra [35], assumes that dh/dt is constant and equal to $-m$:

$$h = -mt + w \ . \tag{2}$$

Here, h is the watertable elevation [L] at a given time; and m [L T^{-1}] and w [L] are obtained by fitting to periods of hydrograph recession.

The power LRC (Figure 2b) adopted from Wendland et al. [34] assumes that dh/dt is a power function of one degree lower ($g-1$) than the recession equation, where $g-1$ is negative:

$$h = f(t-u)^g \,. \tag{3}$$

Here, f [T^{-1}], g [-] and u [T] are fitting parameters.

The exponential LRC (Figure 2c) adopted from Solórzano-Rivas et al. [27] assumes that dh/dt is linearly proportional to h:

$$\frac{dh}{dt} = -ah + b \,, \tag{4}$$

where a [T^{-1}] and b [L T^{-1}] are fitting parameters. Solving for h and applying boundary conditions results in:

$$h\left(t_i\right) = \frac{b}{a} - \left[\left(\frac{-ah_o + b}{a}\right) e^{-a(t_i - t_o)}\right] \,. \tag{5}$$

Here, t_0 [T] and h_0 [L] represent the time and the watertable elevation (respectively) at the beginning of the antecedent recession (point A, Figure 1). The subindex i indicates the timestep.

Equation 2, Equation 3 and Equation 5 were fit to antecedent recession curves using nonlinear regression analysis applying the Python 'curve_fit' function from the SciPy library, version 1.10.1 [31]. The relevant fitting parameters (m and w; f, g and u; or a and b) were obtained by optimising Equation 2, Equation 3 or Equation 5, respectively, against measured heads using the 'Trust Region Reflective' algorithm [31]. The antecedent recession curve was then projected forward to the time of peak h (point C, Figure 1) applying the respective equations, depending on the selected method of extrapolation. Then, Δh_r was calculated as the difference between peak h and the projected watertable elevation (point E, Figure 1), as shown in Figure 2a to Figure 2c. The power LRC solution could not be found for 3 of the 3,000 applications, suggesting possible limitations in its applicability.

The application of MRC approaches included the event-based method described by Nimmo et al. [24] and Nimmo and Perkins [23]. This approach assumes a linear relationship between dh/dt and h (an exponential-decay function), similar to Equation 4, although in this case, a finite-difference approximation of the head gradient $\Delta h/\Delta t (\approx dh/dt)$ was adopted. This is given as:

$$\frac{\Delta h}{\Delta t} = -nh + p \,. \tag{6}$$

Here, n [T^{-1}] and p [L T^{-1}] are fitting parameters. While Equation 4 was integrated to obtain the fitting function for recession extrapolation, the finite-difference alternative given in Equation 6 was applied to individual timesteps to obtain fitting parameters and to project recession curves.

To derive the MRC, recession events were identified as periods of falling groundwater levels (negative $\Delta h/\Delta t$) within the hydrograph record, utilising the 'gradient' algorithm from the NumPy library, version 1.24 [13]. Fitting parameters (n and p) in Equation 6 were then obtained by correlating rates of groundwater decline (negative $\Delta h/\Delta t$) to h, using the curve_fit function (with the Trust Region Reflective algorithm) in the SciPy library [31]. The projected recession curve was then obtained (to the time of peak h) using a daily timestep, as:

$$h_i^* = h_{i-1}^* + \Delta t \left(nh_{i-1}^* + p\right) \,. \tag{7}$$

Here, h^* is the projected watertable elevation [L] and i is the timestep [T]. Δh_r was then calculated as the difference between the hydrograph's peak h (for a given recharge event) and the projected watertable elevation, as shown in Figure 2d. Equation 7 (rather than

Equation 5) was adopted in applying the MRC in the current study to align with the approach of Nimmo et al. [24] and Nimmo and Perkins [23].

A preliminary comparison was undertaken to examine differences between the analytical and finite-difference applications given by Equation 5 and Equation 7, respectively. An individual MRC was constructed (as described in preceding paragraphs) for 1,000 hydrographs (synthetic hydrograph creation is discussed in subsection 2.2). The recession curve was then projected using the fitting parameters obtained from the MRC with both Equation 5 and Equation 7 for the first recharge event of each hydrograph. When the finite-difference approach adopted $\Delta t = 1$ day (as was done in this study), errors in Δh_r (relative to the analytical value) ranged from –0.1% to 0.1%. The difference between the two approaches increased with increasing Δt, as is usual for finite-different approximations. For example, when the finite-difference approach adopted $\Delta t = 30$ days, errors in Δh_r (relative to the analytical value) ranged from –8.1% to 5.0%.

The event-based MRC approach used in this study differs slightly from the approach described by Nimmo et al. [24] and Nimmo and Perkins [23] as their method further reduces Δh_r to account for watertable rise not attributable to recharge, such as changes in temperature, atmospheric pressure and the Lisse effect (described by Crosbie et al. [5] as occurring when intense rainfall traps air in the unsaturated zone causing rapid water level rise). For our application to idealised synthetic hydrographs, reducing Δh_r to account for these phenomena was deemed unnecessary.

The fixed-timestep MRC approach was implemented using the method suggested by Heppner and Nimmo [16]. Firstly, average values of h and $\Delta h/\Delta t$ were calculated for each pair of successive points in the entire hydrograph. Then, following the same approach as the event-based MRC, fitting parameters n and p were obtained by correlating $\Delta h/\Delta t$ recession data to h using Equation 6. The recession curve was subsequently projected in daily timesteps (Figure 2e and Figure 2g) using:

$$h_i^* = h_i - \Delta t\ (p - nh_i)\ . \tag{8}$$

Δh_r was calculated for each timestep as the difference between the observed h_{i+1} and h_i^*. The overall Δh_r for the recharge event was calculated as the sum of values obtained for each timestep. This latter step allowed for a comparison of event-based and fixed-timestep methods, in terms of groundwater discharge estimation.

The RISE approach, as described by Delin et al. [9], was also applied. To estimate recharge for each event, Δh_r was calculated by summing Δh for each daily timestep in the recharge event (Figure 2f and Figure 2h). The bin-averaged methodology of obtaining MRC's [16, 23] was tested, including for both the event-based and fixed-timestep MRC approaches. However, the results obtained were similar to the event-based and fixed-timestep MRC approaches described above, and therefore, the bin-averaged methodology was consequently omitted from this paper for brevity.

The relative error (E_r [%]) in the WTFM-estimated recharge value was calculated for each recharge event using the following equation:

$$E_\mathrm{r} = \left(\frac{R_\mathrm{e} - R_\mathrm{k}}{R_\mathrm{k}}\right) \times 100\ , \tag{9}$$

where R_k is the model input recharge [L T^{-1}].

2.2 Synthetic Hydrograph Creation

One thousand synthetic hydrographs, each with three unique recharge events, were created using Python following the methodology of Solórzano-Rivas et al. [27], who used a form of the

Maasland [22] solution. The solution assumes a homogeneous unconfined aquifer (linearised so that transmissivity (T) [L^2 T^{-1}] is constant) bounded on one side by a no-flow condition, representing a groundwater divide, with a constant-head condition applied to the opposite boundary, representing the effect of a surface feature (e.g., a river or the ocean) receiving groundwater discharge. Intermittent groundwater recharge is uniformly distributed in space. For each hydrograph, the aquifer length (L) [L], distance between the observation well and the groundwater divide (relative to the aquifer length) (ξ) [-], R_k, S_y, and T were randomly selected within pre-defined limits, which are given in Table 2. This was accomplished using the 'random.Generator.uniform' function from the NumPy library [13]. All parameter limits were chosen to match those used by Solórzano-Rivas et al. [27], other than ξ, which adopted the range 0.05 – 0.95 to allow the margins of the aquifer to be analysed. The distance between the observation well and the constant-head boundary (d) [L] is $(1-\xi)L$. The period of recharge was set to 30 days, occurring at times 90 – 120 days, 210 – 240 days and 360 – 390 days, thereby creating periods of hydrograph recession equal to 60, 90 and 120 days, respectively.

Table 2: Parameter limits used to create synthetic hydrographs.

Parameter	Parameter limits	Interval
Transmissivity, T	50 — 3,000 m^2/d	1
Aquifer length, L	1,000 — 5,000 m	1
Specific yield, S_y	0.02 — 0.27 [-]	0.01
Relative observation well location, ξ	0.05 — 0.95 [-]	0.01
Recharge, R_k	0.0005 — 0.0180 m/d	0.0001

3 Results

3.1 Synthetic Hydrograph Application

Figure 3 shows examples of the projected antecedent recession curve for a single recharge event, for which the relative error (E_r) of each WTFM variant was obtained. Negative E_r represents underestimation of R_k, while positive E_r represents overestimation of R_k.

To determine the most accurate WTFM approach, $|E_r|$ was used to rank (from 1 to 6) the WTFM variants for each recharge event. A rank of 1 indicates that the method produced the closest match to the known recharge value (i.e., shown as "0% E_r" in Figure 3), reflecting the highest accuracy, while a rank of 6 indicates the method is furthest from the known recharge value (i.e., highest $|E_r|$), implying lowest accuracy. Figure 4 compares the rankings obtained for the 3,000 simulated recharge events.

Whilst all methods exhibited rankings from 1 to 6, the fixed-timestep MRC approach consistently outperformed others, ranking as 1 for the majority (52%) of recharge events. On average, the fixed-timestep MRC approach underestimated recharge by 10.4% when ranked as 1. The hydrograph model parameters for these events were wide-ranging within the parameter limits listed in Table 2. The fixed-timestep MRC approach ranked as the least accurate (rank 6) for only 1% of the events. For those recharge events, recharge was overestimated by 0.001% to 11.1%. Where this occurred, S_y was greater than 0.1, T was less than 920 m^2/d, and the observation point was located nearer to the groundwater divide ($\xi < 0.63$). The average E_r was positive where the fixed-timestep MRC ranked less than 1, reflecting a tendency to overestimate recharge. These results indicate that where the fixed-timestep MRC approach

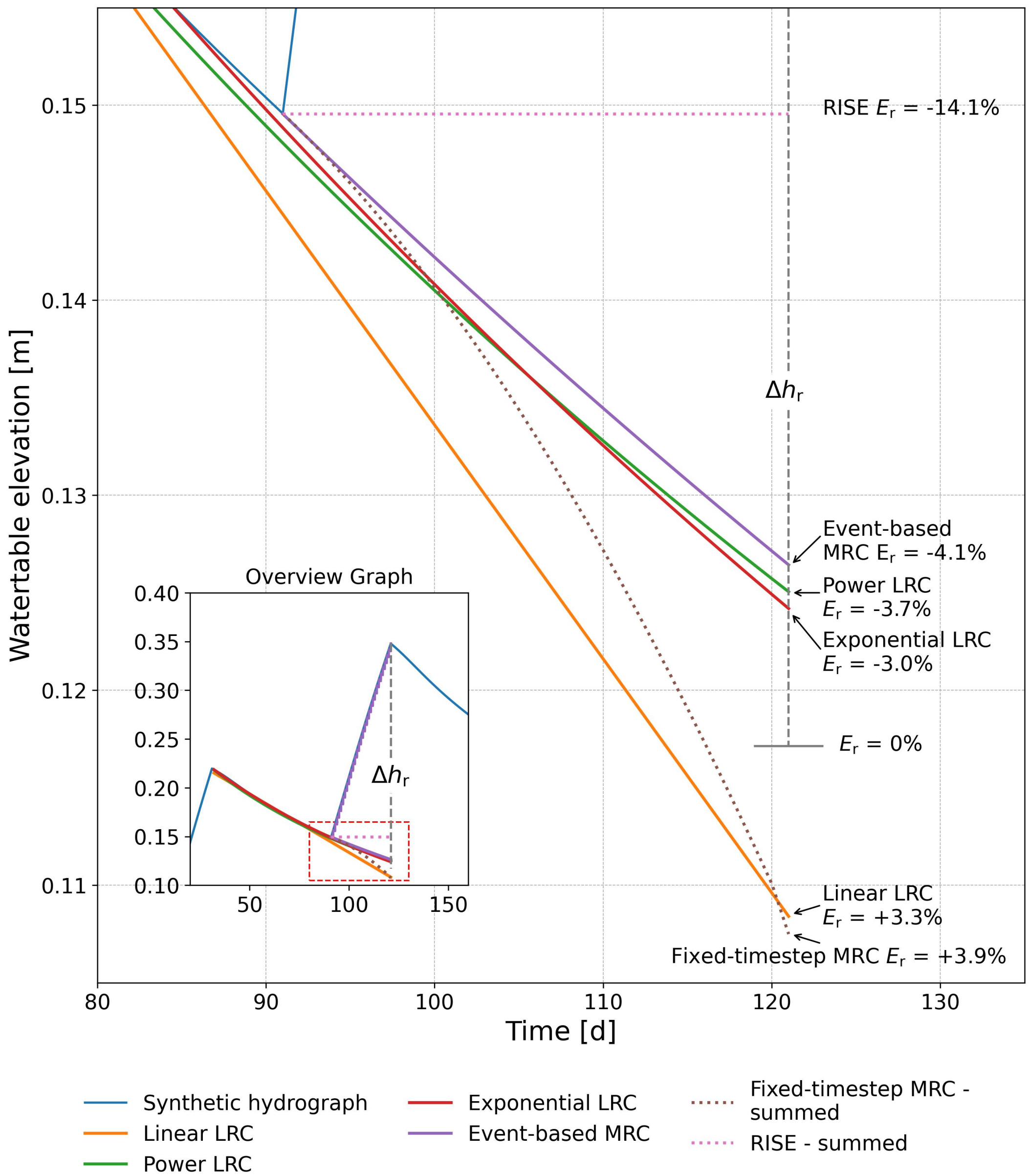


Figure 3: Example of relative error (E_r) values obtained for each WTFM variant when applied to a synthetic hydrograph, where $R_k = 0.002$ m/d, $T = 2450$ m^2/d, $L = 2230$ m, $S_y = 0.26$ and $\xi = 0.5$. The Δh_r required to estimate recharge correctly is labelled as $E_r = 0\%$. Summed values from the RISE and fixed-timestep MRC approaches (shown as dotted lines) represent the summation of projected recession over individual timesteps during the period of watertable rise.

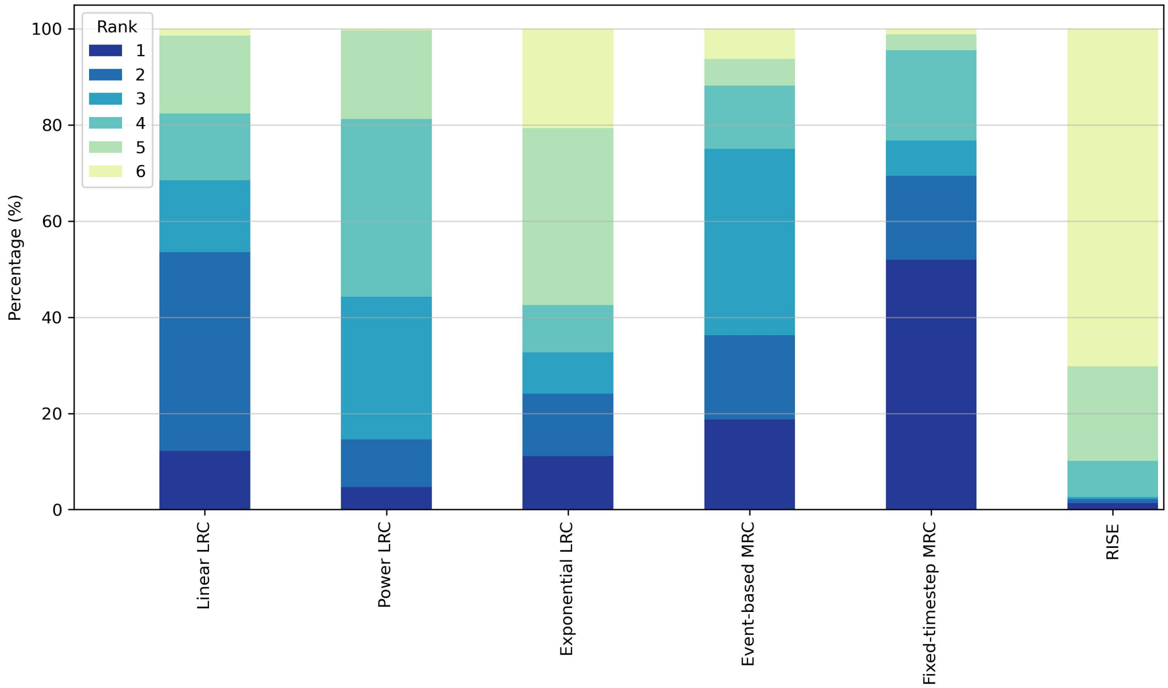


Figure 4: WTFM variants ranked from 1 to 6 after application to 3,000 individual recharge events. A rank of 1 identifies WTFM applications that produced the lowest error (smallest $|E_\mathrm{r}|$), while a rank of 6 represents the method with the greatest $|E_\mathrm{r}|$ for each recharge event. The percentage refers to the proportion of recharge events assigned to a given ranking.

is the best choice, it tends to underestimate recharge, but conversely, where it overestimates recharge, other WTFM approaches tend to produce an improved value of recharge.

The event-based MRC approach ranked as 1 for the second-highest portion of events (18.7%) with an average E_r of -1% for these cases. This was followed by the linear LRC approach, which ranked as 1 for 12.2% of events with an average E_r of -0.3% for these cases. Conversely, the event-based MRC approach ranked as 6 more often than the linear LRC approach (6.2% and 1.4% of cases, respectively), indicating that the linear LRC approach and the fixed-timestep MRC approach are the least likely to produce the poorest predictions of recharge, amongst the methods tested. The RISE method proved the least accurate of the six variants, ranking as 6 for 70.2% of the recharge events with an average E_r of -17.6% for these cases. This reflects the need to incorporate estimates of groundwater discharge into WTFM approaches given the lack of discharge effects in the RISE method. Table A1 lists error statistics for each WTFM variant and for each ranking, including the range of E_r, the average E_r, the average $|E_\mathrm{r}|$, and the count.

Figure 5 presents box plots of E_r for the six WTFM variants, showing that the linear LRC, exponential LRC and fixed-timestep MRC had median E_r values close to 0%. However, the interquartile range (i.e., the middle 50% of results) was much narrower for the linear LRC and fixed-timestep MRC than for the exponential LRC, confirming that the linear LRC and fixed-timestep MRC estimated recharge more accurately for a wider range of recharge rates and aquifer parameters. The power LRC, exponential LRC and event-based MRC tended to underestimate recharge across the range of conditions assessed. We attribute this to the decreasing rate of groundwater discharge that is assumed in those methods as the watertable rises, as explained in section 1. The median, minimum, maximum and the 25$^{\mathrm{th}}$ and 75$^{\mathrm{th}}$ percentile values that define the box plots in Figure 5 are provided in Table A2.

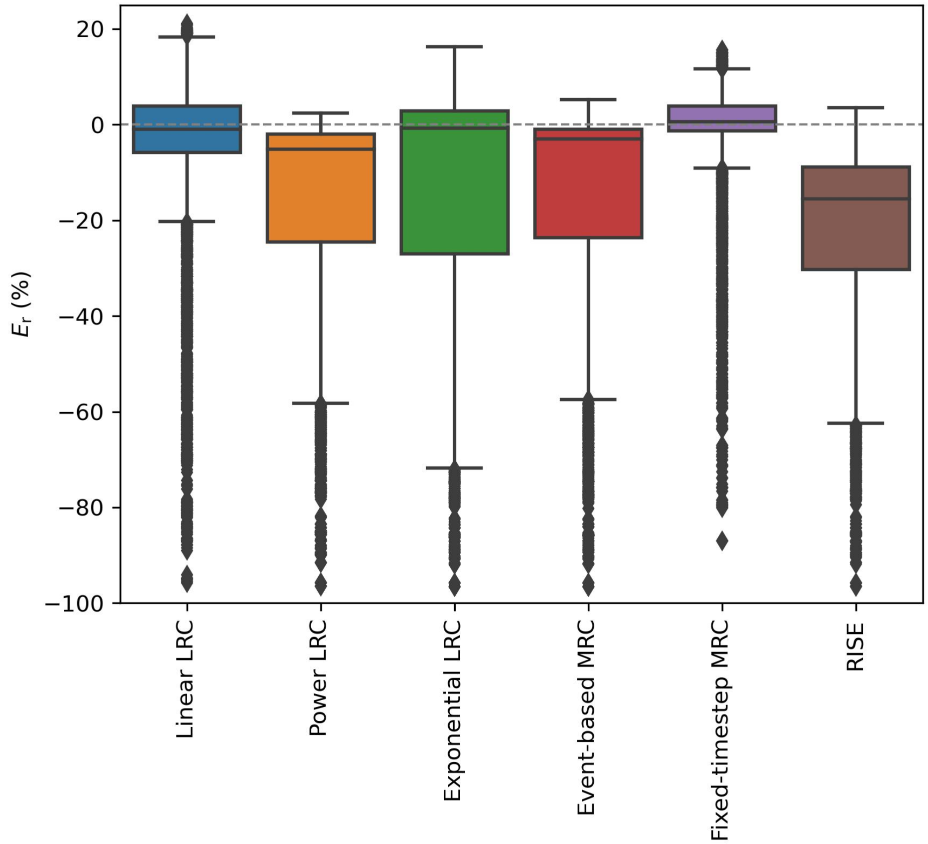


Figure 5: Box plot illustrating the distribution of E_{r} from the application of six variants of the WTFM. The median is represented by the line within each box, while the upper and lower edges of the box correspond to the 25$^{\mathrm{th}}$ and 75$^{\mathrm{th}}$ percentiles (i.e., the *interquartile range*), respectively. Whiskers extend to 1.5 times the interquartile range from the quartiles (i.e., edges of the box), with larger error values represented as individual points. The grey dashed line indicates 0% E_{r}.

All variants of the WTFM exhibited large error values (shown as individual points in Figure 5), with the vast majority being underestimates of recharge, with E_{r} approaching -100% in extreme cases. Only the linear LRC and fixed-timestep MRC exhibited large errors that overestimated recharge. Figure 6 explores relationships between the occurrence of large errors and model input parameters for the fixed-timestep MRC approach, showing E_{r} in relation to ξ and third variables of T, R_{k}, L and S_{y}. Figure 7 presents the same analysis for the exponential LRC approach. Similar plots for the remaining WTFM variants are available in Appendix A (Figure A1 to Figure A4).

Figure 6 shows that, for the fixed-timestep MRC approach, large overestimation of recharge ($E_{\mathrm{r}} > 12\%$) occurred at observation points relatively close to the groundwater divide ($0.05 < \xi < 0.22$). For the 64 cases where this occurred, the mean L (1,639 m) was 55% of the overall mean L of 2,976 m, calculated from all 3,000 cases. No clear relationship was observed between the occurrence of large overestimation errors and model input parameters T, R_{k} and S_{y}. Large overestimation errors ($E_{\mathrm{r}} > 18\%$) were also observed for the linear LRC method (in 28 cases), again at observation points near the groundwater divide ($0.05 < \xi < 0.16$; Figure A1). The mean L (1,636 m) of these cases was also 55% of the overall mean L, with no clear relationship observed between occurrences of large overestimation errors and model input parameters T, R_{k} and S_{y}. These results suggest that ξ and L are key factors influencing the large overestimation

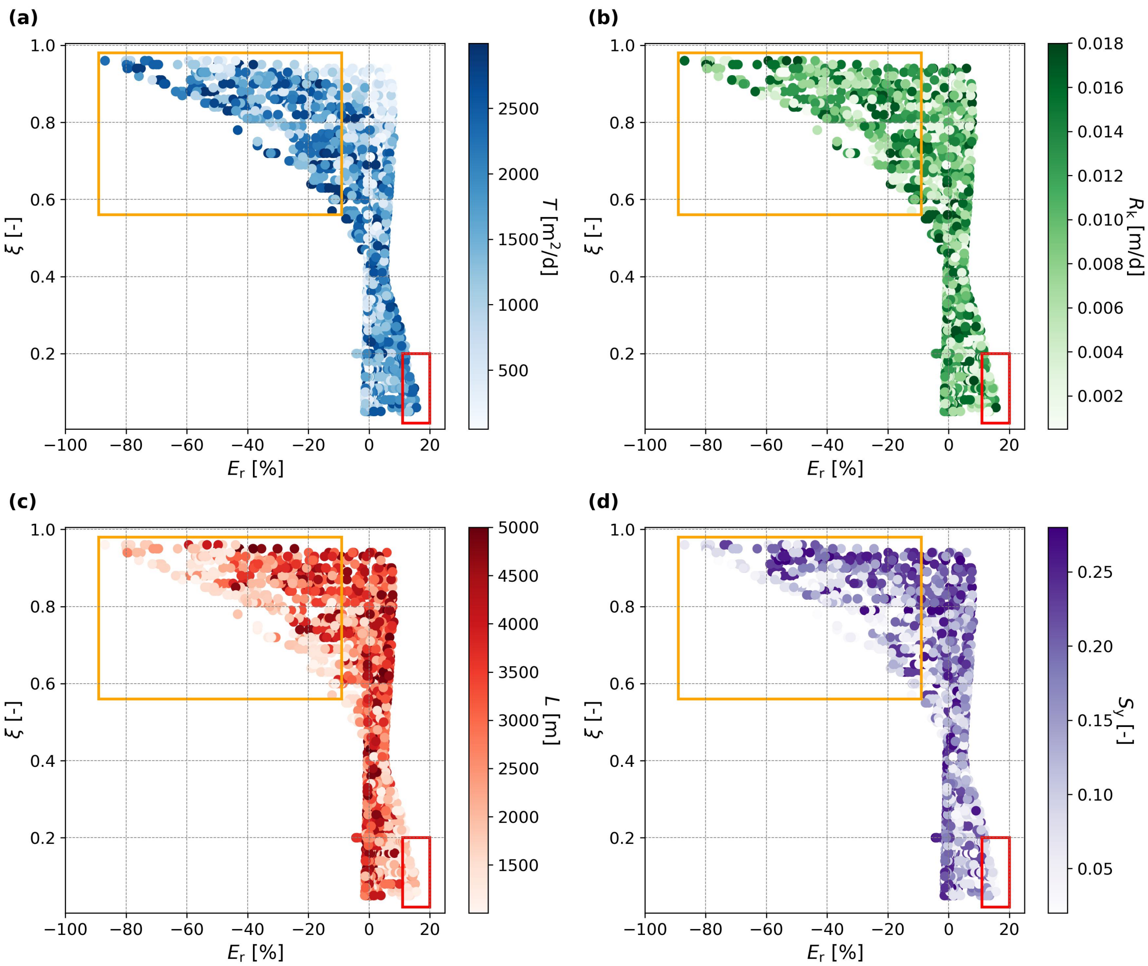


Figure 6: Scatter plots illustrating relationships between E_{r} and ξ for the fixed-timestep MRC approach: (a) E_{r} vs ξ with respect to T, (b) E_{r} vs ξ with respect to R_{k}, (c) E_{r} vs ξ with respect to L, (d) E_{r} vs ξ with respect to S_{y}. The orange and red rectangles highlight large error values (represented as individual points in Figure 5) that underestimate and overestimate recharge, respectively.

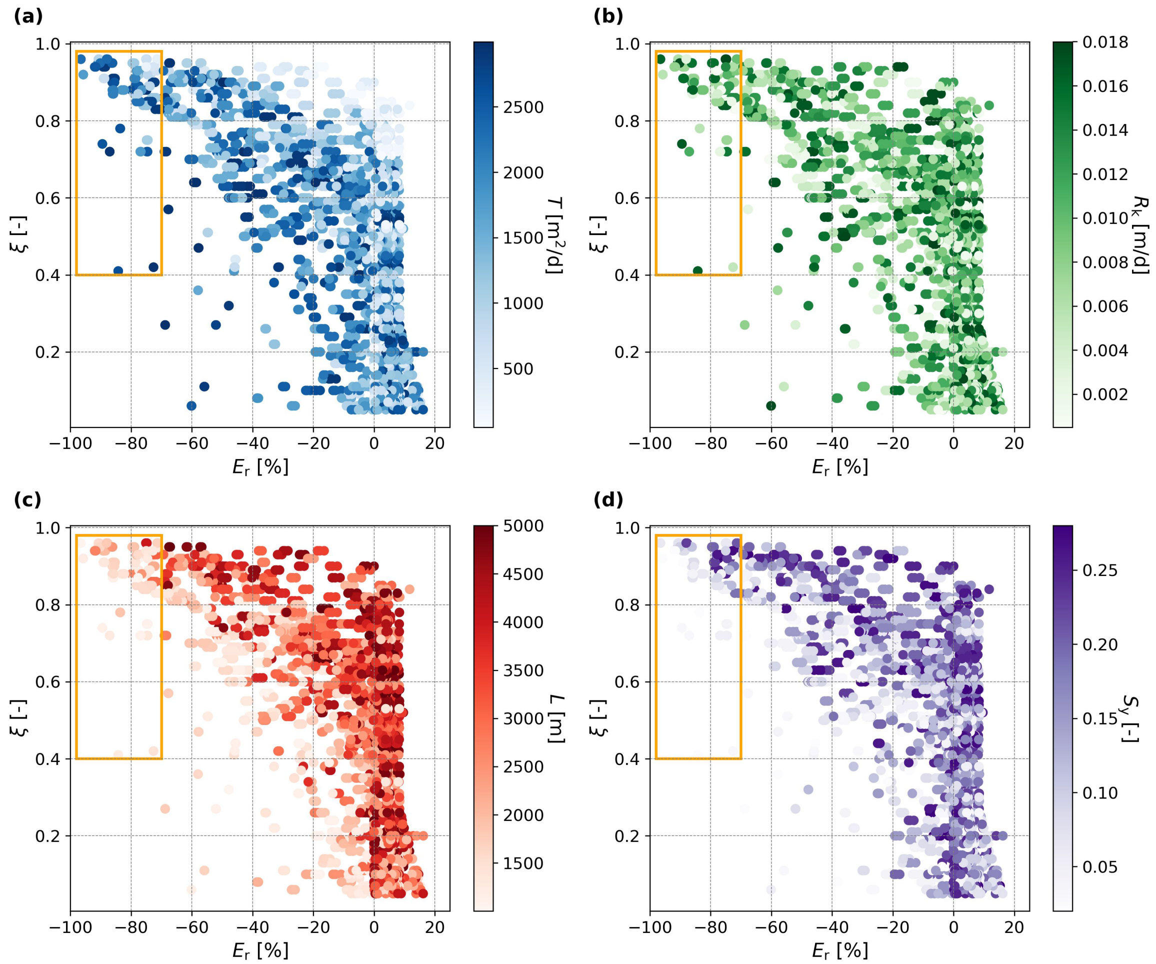


Figure 7: Scatter plots illustrating relationships between E_r and ξ for the exponential LRC approach: (a) E_r vs ξ with respect to T, (b) E_r vs ξ with respect to R_k, (c) E_r vs ξ with respect to L, (d) E_r vs ξ with respect to S_y. The orange rectangle highlights large error values (represented as individual points in Figure 5) that underestimate recharge.

of recharge by the fixed-timestep and linear LRC approaches. Table A3 provides the mean, standard deviation, minimum and maximum values for model input parameters, including for cases where a large overestimation of recharge occurred (for the fixed-timestep and linear LRC methods) and for the overall dataset.

While the exponential LRC, power LRC, event-based MRC and RISE methods did not produce *large* overestimation errors (as defined above), they did overestimate recharge for 1,411, 159, 329 and 30 of the 3,000 cases, respectively. For the exponential LRC approach, the 25 cases that produced the maximum overestimation of recharge ($E_{\mathrm{r}} > 12\%$) occurred relatively close to the groundwater divide ($0.05 < \xi < 0.2$; Figure 7), as occurred in applications of the fixed-timestep MRC and linear LRC approaches. Conversely, for the power LRC and event-based MRC, the maximum recharge overestimation occurred closer to the constant-head boundary ($0.67 < \xi < 0.9$; Figure A2 and Figure A3), albeit the number of cases in which this occurred was small for those approaches. No trend was observed between ξ and the overestimation of recharge for the RISE method. Taking the various methods together, it appears that application of the WTFM in observation wells relatively close to groundwater divides has a greater potential to produce recharge overestimation in all approaches aside from the RISE method. For some methods (power LRC and event-based MRC), observation wells relatively close to fixed-head boundaries (locations of groundwater discharge) may also produce recharge overestimation.

The linear LRC, power LRC and event-based MRC variants produced large underestimation errors across the entire range of ξ tested, as evident in Figure A1, Figure A2 and Figure A3, respectively. Large underestimation errors from the fixed-timestep MRC approach occurred more so when the observation well was sited closer to the constant-head (groundwater outflow) boundary ($0.55 < \xi < 0.95$; Figure 6). This was also apparent from the results of the exponential LRC and RISE methods, although for larger ranges in ξ ($\xi > 0.4$ and $\xi > 0.27$, respectively; Figure 7 and Figure A4). For all WTFM variants tested, more than 90% of large underestimation cases occurred when ξ was greater than 0.6. For the exponential LRC, RISE, linear LRC, power LRC and event-based MRC methods, where large underestimation errors occurred at a ξ of less than 0.6, L was less than 1,759 m and S_{y} was smaller than 0.12. Lower values of S_{y} and L did not appear to increase the occurrence of large underestimation errors for the fixed-timestep MRC approach when the observation well was located further from the constant-head boundary ($\xi < 0.6$) (Figure 6c and Figure 6d).

The widest range of E_{r} was found to occur relatively close to the constant-head boundary for all WTFM variants tested, and this appeared to be unaffected by R_{k} or S_{y}. However, higher values of T ($> 1{,}500$ m^2/d) were associated with greater recharge underestimation for cases where $0.8 < \xi < 0.95$ (near the constant-head boundary), while E_{r} ranged from overestimation to large underestimation when $T < 1{,}500$ m^2/d (e.g., Figure 7a).

Smaller values of L ($< 2{,}000$ m) appeared to lead to greater recharge underestimation for all WTFM variants where the observation well was relatively close to the constant-head boundary ($0.8 < \xi < 0.95$), whereas when L exceeded 2,000 m, both overestimation and large underestimation of recharge occurred (e.g., Figure 7c). The role of L on the under/overestimation of WTFM recharge was explored further through sensitivity analysis. Additional scenarios were undertaken with L fixed (for all realisations), at values of 1,000 m and 5,000 m. This allowed for the effects of L on the WTFM error to be distinguished from the effect of d (distance to the constant-head boundary). The relative position of the observation well (i.e., ξ) was set so that values of d were comparable in both the 1,000 m cases and the 5,000 m cases. The results, provided in Figure A5, indicate that for scenarios where $L = 5{,}000$ m, and d ranged between 50 m and 250 m, recharge was underestimated by at least 35%, 48%, 51%, 19% and 19% for the linear LRC, power LRC, exponential LRC, event-based MRC, fixed-timestep MRC and RISE method, respectively. Whereas, when d was 250 m to 500 m (for $L = 5{,}000$ m), E_{r}

ranged from overestimation of recharge to large underestimation of recharge for all WTFM variants, except for the RISE method, for which recharge was underestimated by 0.7% to 87%. This trend of increasing underestimation of recharge for cases with smaller d (observation well closer to the constant-head boundary) in the $L = 5{,}000$ m cases also occurred for the scenarios where $L = 1{,}000$ m. Thus, we conclude that d and T are the primary controls on the magnitude of recharge underestimation near the constant-head boundary, rather than L. The initial indication that recharge underestimation is worse for small values of L arose because the observation point was more likely to occur closer to the constant-head boundary (i.e., d was smaller on average) for cases of smaller L. d was not found to impact results near the groundwater divide.

3.2 Field Data Application

To evaluate the practical challenges associated with implementing the WTFM and to validate the findings of the synthetic study, we applied each WTFM variant to six recharge events observed in observation well RN12030004. This well is situated within an unconfined alluvial aquifer in North Queensland, Australia (Latitude -20.4844, Longitude 145.4822). The bore casing is screened between 13.35 to 15.35 m below ground level, intersecting a stratigraphic sequence of coarse and fine sands, within which the watertable fluctuates.

For the fixed-timestep and event-based MRC approaches, the MRC fitting parameters were obtained using MRCfit [23], yielding n = 0.00106 and p = -0.00201. MRCfit allowed hydrograph recession periods that occurred within 5 days of rainfall events to be excluded from the MRC derivation, which improved the robustness of the fit. Such exclusions were not required in the analyses of synthetic hydrographs. Recharge estimates were subsequently calculated using the methodology described in subsection 2.1.

Figure 8 presents a graphical comparison of the WTFM variants applied to observed recharge events at RN12030004. Figure 8a shows the observation record of daily water levels, with coloured overlays indicating the timing of the six recharge events analysed. Figure 8b to Figure 8e illustrate the application of the WTFM variants to four selected recharge events. Table 3 summarises the recharge estimates for all six watertable-rise events assuming S_y = 0.1, which is the minimum S_y for a fine sand [19].

Table 3: Recharge estimates (mm) for each watertable-rise sojourn in RN12030004 (between 2018 and 2021) assessed using six WTFM variants. S_y is assumed to equal 0.1.

	Recharge/event (mm)					
WTFM variant	**2018**	**2019a**	**2019b**	**2020**	**2021a**	**2021b**
Linear LRC	264	443	188	221	335	94
Power LRC	217	NA	186	204	NA	93
Exponential LRC	207	433	237	215	303	106
Event-based MRC	255	413	161	190	314	91
Fixed-timestep MRC	261	432	164	196	325	91
RISE	218	375	107	144	270	65

Of the six recharge events analysed, the RISE approach produced the lowest recharge estimate for five. The exception was the 2018 recharge event (Figure 8b), where the exponential LRC approach yielded the lowest recharge estimate. This is similar to the result of the theoretical analysis, in which the RISE method estimated the least recharge for 80% of

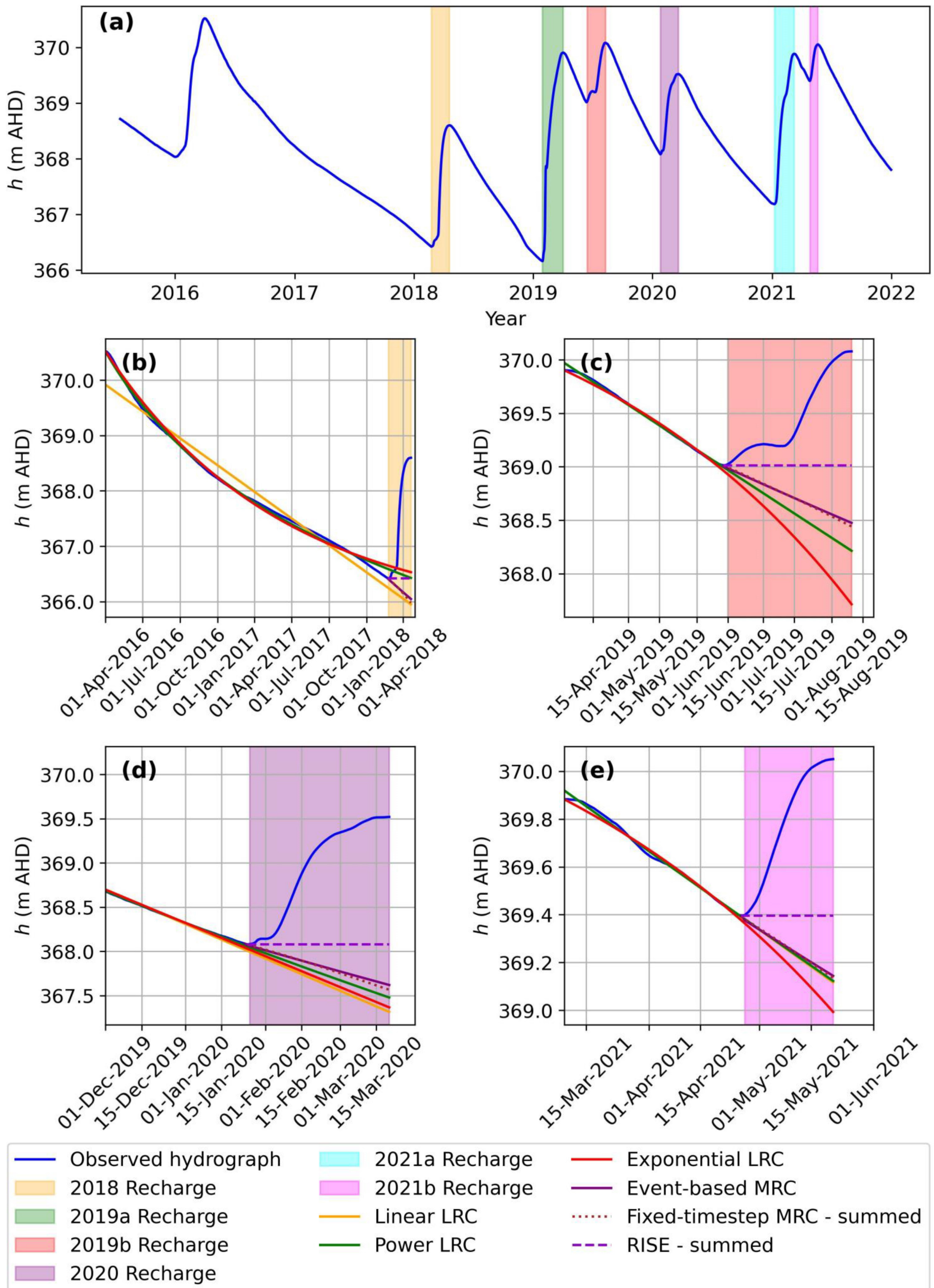


Figure 8: Graphical comparison of the WTFM approaches applied to field data at site RN12030004: (a) Daily water level observations; WTFM variants applied to the: (b) 2018 recharge event, (c) 2019b recharge event, (d) 2020 recharge event, and (e) 2021b recharge event. Shaded regions highlight the recharge events analysed in this study. Descriptions of the respective methods are given in subsection 2.1.

the recharge events, while the exponential LRC approach did so for 12% of recharge events. Underestimation of the 2018 recharge event occurred with the LRC approach even though an excellent fit was obtained between the antecedent recession curve and both the exponential and power functions ($r^2 > 0.995$), and so a poor fit to the recession period was not the cause. Rather, the projected recession curves of both the exponential and power functions deviated from the observed hydrograph prior to the lowest point in the recession, and this appears to have contributed to the underestimation of recharge. Given that the RISE method consistently underestimated the rate of recharge in the prior analysis of synthetic hydrographs, it appears that the exponential (and power) LRC approach underestimated recharge in 2018 given that it was exceeded by the RISE method.

The linear LRC approach provided a poorer fit ($r^2 = 0.967$) to the antecedent recession curve relative to the exponential and power functions, yet its recharge estimate was within 9 mm of those obtained using the event-based MRC and fixed-timestep MRC approach (for the 2018 recharge event). Based on the findings of the prior analysis of synthetic hydrographs, we expect that the latter estimates are the most reliable of the methods tested, at least for the 2018 event.

The linear LRC approach yielded the maximum recharge estimates for four of the six events. For the 2019b and 2021b recharge events, both characterised by shorter antecedent recession periods (72 and 48 days respectively), the exponential LRC approach produced the maximum recharge estimates. The theoretical analysis indicates similar results, with the linear LRC and exponential LRC approaches estimating the maximum recharge for 26% and 36% of the synthetic recharge events, respectively. The fixed-timestep MRC approach obtained the maximum recharge for 38% of the synthetic recharge events. For the 2019b and 2021b recharge events, the exponential LRC approach fitted a downward-sloping curve to the antecedent recession (Figure 8c and Figure 8e). Downward-sloping curves were also found to occur in the theoretical analysis. Solórzano-Rivas et al. [27] found that when the exponential LRC approach projected a downward sloping curve, recharge was overestimated, and we expect that this also occurred here, notwithstanding that the "true" recharge cannot be known from the field dataset.

The event-based MRC, fixed-timestep MRC and RISE approaches, which all commence projection at the lowest point of the recession (point B, Figure 1), tended to yield results in a consistent order for each recharge event when applied to the observed data. Here, the fixed-timestep MRC approach estimated more recharge than the event-based MRC approach, which estimated more recharge than the RISE method for all six recharge events. This is similar to the synthetic hydrograph analysis, where the fixed-timestep MRC estimated more recharge than the event-based MRC approach for 99.8% of cases. In the synthetic analysis, the event-based MRC method produced more recharge than the RISE method for 91% of the events. For the remaining 9% of cases, the event-based MRC method extrapolated the recession (to point E; Figure 1) using a curve that increased in time (rather than a negative slope, as expected from the extrapolation of a recession curve), thereby leading to a lower recharge value than that produced by the RISE method. The positive slope produced by recession extrapolation using the exponential function in the event-based MRC approach arose because it had an asymptote (p/n) at a higher elevation than the lowest point of the recession (point B, Figure 1) being assessed.

Overall, the application of the WTFM variants to a real-world hydrograph reinforces the findings of the synthetic hydrograph analysis. For example, large variability in recharge estimates was produced by the exponential LRC approach when applied to both synthetic hydrographs and the field data, where it generated both the maximum and minimum recharge estimates for different events. The power LRC approach, which showed less variability in recharge estimates compared to the exponential LRC method, failed to generate a solution for three recharge events in the synthetic study and for two recharge events in the field application, highlighting that this method can occasionally become unstable. The fixed-timestep and event-based variants of the MRC approach estimated recharge consistently, at least relative to each other, for both the field data and synthetic hydrographs.

Applying the WTFM variants to observed hydrographs introduces practical challenges that were not faced in the synthetic study. For instance, developing the MRC requires careful selection of recession periods that represent the 'natural discharge' from the aquifer. Identifying the start and end of watertable rises attributable to recharge is no menial task, along with determining if the observed water level rise required some level of subjective decision making to assess whether the observed water level rise is solely due to recharge or influenced by other factors such as air entrapment. Additionally, as highlighted by Crosbie et al. [5], selecting an appropriate S_y is critical for accurate recharge estimates, although this aspect was beyond the scope of this study.

4 Discussion

This study evaluated alternative forms of the WTFM to determine the most accurate approach to estimate Δh_r. The fixed-timestep MRC method proved to be the most consistently accurate, followed by (in descending order of accuracy) the linear LRC approach, the event-based MRC approach, the power and exponential LRC approaches, and lastly the RISE method. Despite its accuracy, the fixed-timestep MRC method seems to be underutilised in practice. For example, a recent review by Becke et al. [3] of 40 published WTFM applications showed that only seven studies used the fixed-timestep MRC approach. Of the 40 studies, 24 adopted recession-extrapolation methods (the exponential LRC, power LRC and event-based MRC approaches) that assume that discharge decreases as the watertable rises (i.e., the opposite of Darcy-based flow principles, as discussion in section 1), and therefore, we expect those investigations to have under-estimated the watertable change attributable to recharge. Whether the recharge was also underestimated depends on the accuracy of specific yield values. The underutilisation of fixed-timestep MRC methods may stem from the need to have water levels at a high temporal resolution, such as daily measurements, which are commonly not available. Nevertheless, the use of WTFM approaches that presume declining discharge as the watertable rises ought to be reconsidered given the results of the current study.

For the majority of cases tested, all WTFM variants tended to underestimate recharge. This is consistent with the findings of Solórzano-Rivas et al. [27], who also tested the event-based MRC and exponential LRC approaches on synthetic hydrographs. Conversely, recharge overestimation was also found to occur, particularly when applying the linear LRC, exponential LRC or fixed-timestep MRC approaches to observation points near the groundwater divide.

Field studies using multiple recharge estimation techniques report various outcomes regarding the accuracy of recharge estimation using the WTFM. For example, Hagedorn et al. [12], Hung Vu and Merkel [17] and Somaratne et al. [28] found the WTFM to be the most accurate approach when estimating recharge in Jeju Island (Korea), Hanoi (Vietnam) and South Australia (Australia), respectively. However, Barua et al. [2], Cartwright et al. [4], King et al. [20], von Freyberg et al. [32] and Yenehun et al. [36] concluded that WTFM approaches overestimated recharge when compared to other field-based recharge estimation techniques. The latter included chloride mass balance, water balance methods and unsaturated zone modelling. The apparent overestimation of WTFM recharge when the method is applied to real-world data may arise from multiple causes, some of which have not been captured when analysing the approach on synthetic hydrographs as we have undertaken. For example, 'noise' in real-world hydrographs (i.e., watertable fluctuations due to causes unrelated to recharge, such as atmospheric pressure changes, pumping, evapotranspiration, etc.; [23]) may be interpreted as recharge in applications of the WTFM. It is likely that fixed-timestep methods are more susceptible (than event-based methods) to the incorrect interpretation of noise as recharge events, because event-based methods assess recharge over longer timeframes, and therefore tend to neglect small watertable variations. Crosbie et al. [5] demonstrated that filtering high-temporal resolution water-level data to remove noise and non-rainfall-related watertable fluctuations prior to applying the fixed-timestep MRC approach can mitigate these effects.

The choice of S_y is also known to result in significant uncertainty in WTFM recharge estimates. For example, Crosbie et al. [5] found that using a constant S_y, particularly in aquifers with a shallow unsaturated zone, could produce recharge estimates up to 70% higher than those derived using a depth-dependent S_y. In some cases, this overestimation of recharge may outweigh the underestimation caused by incorrectly characterising discharge (i.e., the main source of error investigated in the current study), potentially leading to an overall overestimation of recharge if S_y is significantly overestimated (as occurs when the watertable reaches the land surface; [5]).

Our findings show that the largest underestimation of recharge occurred near the constant-head boundary (i.e., small values of d or large values of ξ) for all WTFM variants tested, which is consistent with the conclusions for the RISE method reported by Águila et al. [1]. However, for the exponential LRC, RISE, linear LRC, power LRC and event-based MRC methods, we observed that the combination of relatively low values of S_y and L can result in large underestimation of recharge across a wide range of ξ values. However, low values of S_y and L did not produce large underestimation of recharge when applying the fixed-timestep MRC approach, further underscoring the potential benefits of adopting this method.

The impact of R_k (model input recharge) on recharge estimation error in the current study differed from prior studies. Águila et al. [1] reported that errors made in estimating recharge with the RISE method increased with the rate of recharge (for uniform recharge over 30 days, which was also used in the current study). In contrast, our study found no relationship between R_k and E_r for any of the methods tested. This discrepancy is potentially due to our study testing the rate of recharge in a wider variety of scenarios, or alternatively it could be due to our study testing a maximum recharge rate of 547 mm/month compared to 1,000 mm/month by Águila et al. [1].

5 Conclusions

Of the watertable fluctuation method (WTFM) variants tested, the fixed-timestep master recession curve (MRC) approach proved to be the most consistently accurate. This finding is attributed to its head-dependent extrapolation of the recession curve, which assumes that discharge increases as the watertable rises, consistent with Darcy's Law for groundwater discharge to surface water bodies of relatively stable water levels. Despite its higher accuracy, the fixed-timestep MRC approach appears to be underutilised in practice, according to a prior review of existing WTFM applications by Becke et al. [3], with other approaches such as the exponential LRC, power LRC and event-based MRC more often employed. The linear local recession curve (LRC) method, which assumes a constant rate of discharge as the watertable rises, was the second most accurate variant. We attribute the lower accuracy of other approaches, such as the exponential LRC, power LRC and event-based MRC to the lower discharge rates that are adopted as the watertable elevation rises. The RISE method, which does not account for ongoing discharge, was the poorest performing method tested.

The widest range of recharge estimation error, for all WTFM variants tested, were observed relatively near the constant-head boundary, where recharge estimation error ranged from slight overestimation to large underestimation. Near the constant-head boundary, cases with low d (distance between the observation well and the constant-head boundary) and high T (transmissivity) produced the largest underestimation of recharge. Cases with low L (aquifer length) and low S_y (specific yield) can also lead to large underestimation of recharge over a wide range of ξ (distance between the observation well and the groundwater divide (relative to the aquifer length)) for all WTFM variants tested other than the fixed-timestep MRC approach. Overestimation of recharge was more pronounced closer to the groundwater divide when applying the linear LRC, fixed-timestep MRC and exponential LRC approaches, with low L exacerbating the errors for the linear LRC and fixed-timestep MRC approaches.

These results challenge the long-held assumption that the projected recession curve should represent the decline in the watertable that would have occurred in the absence of recharge, typically represented by exponential decay, causing the assumed rate of groundwater discharge in WTFM methods to decrease as the watertable rises. Instead, we demonstrate that utilising a recession curve that becomes progressively steeper as the watertable rises improves recharge estimation accuracy. This finding has practical implications for sustainable groundwater management of water resources, including the protection of ecosystems and the prediction of system reliability under increasing water demands.

This study further characterises the understanding of recharge estimation errors when applying the WTFM to homogeneous aquifers. Future research efforts should focus on assessing the reliability of WTFM recharge estimates in more complex aquifer settings, such as those with heterogeneity, fluctuating head boundaries and temporally and spatially varying recharge. Addressing these challenges is essential for improving the accuracy and reliability of groundwater recharge estimates obtained using WTFM techniques.

Acknowledgements

This research was supported by Lower Burdekin Water, whose funding and assistance were instrumental to the project. This research was also partially funded by the Australian Government through the Australian Research Council (project code LP210100430). Amy L. Becke acknowledges the support of an Australian Government Research Training Program scholarship. We also extend our gratitude to the anonymous reviewer and editor for their valuable feedback and contributions.

Data Availability

Python code for this paper is available at github.com/beckea/WTFM_Comparison.

Author Contributions

ALB: Writing - original draft, Investigation, Visualisation.
SCSR: Conceptualisation, Writing - review & editing, Supervision.
ADW: Writing - review & editing, Supervision, Funding acquisition, Project administration.

Appendix A

Table A1: The range of E_r; the average E_r and average $|E_r|$ (presented in brackets); and the count (presented in italics) of each WTFM variant for each rank.

Rank	**Linear LRC**	**Power LRC**	**Exponential LRC**	**Event-based MRC**	**Fixed-timestep MRC**	**RISE**
1	-26.5 to 12.4% (-0.3%, |3.2%|) *366*	-3.3 to 2.5% (-0.2%, |0.7%|) *139*	-12.7 to 2.5% (-0.9%, |1.4%|) *333*	-12.9 to 3.1% (-1.0%, |1.4%|) *562*	-87.0 to 15.6% (-10.4%, |12.0%|) *1560*	-0.5 to 0.0% (-0.17%, |0.17%|) *40*
2	-95.7 to 14.6% (-21.4%, |23.1%|) *1240*	-7.2 to 2.1% (-0.9%, |1.5%|) *298*	-14.2 to 4.9% (-0.3%, |2.2%|) *389*	-13.9 to 4.0% (-2.2%, |2.4%|) *525*	-26.6 to 15.2% (1.3%, |4.3%|) *523*	-0.7 to 0.6% (-0.3%, |0.4%|) *25*
3	-6.1 to 13.9% (-0.7%, |3.0%|) *450*	-96.5 to 1.0% (-26.5%, |26.5%|) *890*	-29.2 to 6.3% (-1.7%, |4.5%|) *260*	-89.4 to 4.7% (-17.3%, |17.6%|) *1165*	0.0 to 13.8% (4.3%, |4.3%|) *221*	-96.5 to -0.0% (-52.4%, |52.4%|) *14*
4	-5.4 to 15.9% (-0.2%, |2.6%|) *416*	-96.5 to 0.0% (-18.6%, |18.6%|) *1109*	-68.8 to 7.3% (-12.7%, |15.7%|) *295*	-91.7 to 0.3% (-29.0%, |28.9%|) *394*	-0.0 to 13.7% (4.8%, |4.8%|) *561*	-96.5 to 3.5% (-51.0%, |51.1%|) *225*
5	-5.1 to 19.3% (7.8%, |7.8%|) *486*	-89.4 to 0.0% (-9.3%, |9.3%|) *555*	-96.5 to 12.9% (-18.6%, |22.3%|) *1104*	-96.5 to 0.5% (-27.1%, |27.1%|) *166*	-0.0 to 10.5% (3.3%, |3.3%|) *100*	-91.8 to 0.34% (-30.9%, |30.7%|) *589*
6	15.7 to 21.1% (18.4%, |18.4%|) *42*	-68.8 to -0.0% (-14.5%, |14.5%|) *9*	-96.6 to 16.2% (-29.3, |35.2%|) *619*	-96.5 to 5.2% (-47.0%, |47.1%|) *188*	0.0 to 11.1% (0.42%, |0.42%|) *35*	-72.6 to 0.4% (-17.6%, |17.6%|) *2107*

Table A2: Descriptive statistics of the E_r for each WTFM variant.

	Linear LRC	**Power LRC**	**Exponential LRC**	**Event-based MRC**	**Fixed-timestep MRC**	**RISE**
Median	-0.98	-5.18	-0.69	-2.94	0.58	-15.52
Interquartile Range	9.67	22.50	29.88	22.58	5.18	21.45
Minimum	-95.74	-96.51	-96.57	-96.55	-86.95	-96.52
Maximum	21.09	2.46	16.22	5.20	15.58	3.54
25th percentile	-5.80	-24.50	-27.00	-23.58	-1.30	-30.31
75th percentile	3.88	-2.00	2.88	-1.00	3.88	-8.86

Table A3: Mean, standard deviation, minimum and maximum of model input parameters for the 64 fixed-timestep MRC and 28 linear LRC cases where recharge was largely overestimated, and for the entire population (3000 cases).

Model input parameter	**WTFM variant**	**Mean**	**Standard deviation**	**Minimum**	**Maximum**
T	Fixed-timestep MRC	1940	632	628	2906
	Linear LRC	1600	494	628	2681
	Entire population	1496	853	50	2998
ξ	Fixed-timestep MRC	0.11	0.04	0.05	0.22
	Linear LRC	0.09	0.03	0.05	0.16
	Entire population	0.50	0.26	0.05	0.95
L	Fixed-timestep MRC	1639	468	1061	2988
	Linear LRC	1636	622	1143	3724
	Entire population	2976	1151	1001	5000
R_k	Fixed-timestep MRC	0.0086	0.0045	0.0007	0.0174
	Linear LRC	0.0076	0.0038	0.0007	0.0151
	Entire population	0.0093	0.0051	0.0005	0.0180
S_y	Fixed-timestep MRC	0.08	0.06	0.02	0.24
	Linear LRC	0.12	0.07	0.02	0.25
	Entire population	0.15	0.08	0.02	0.27

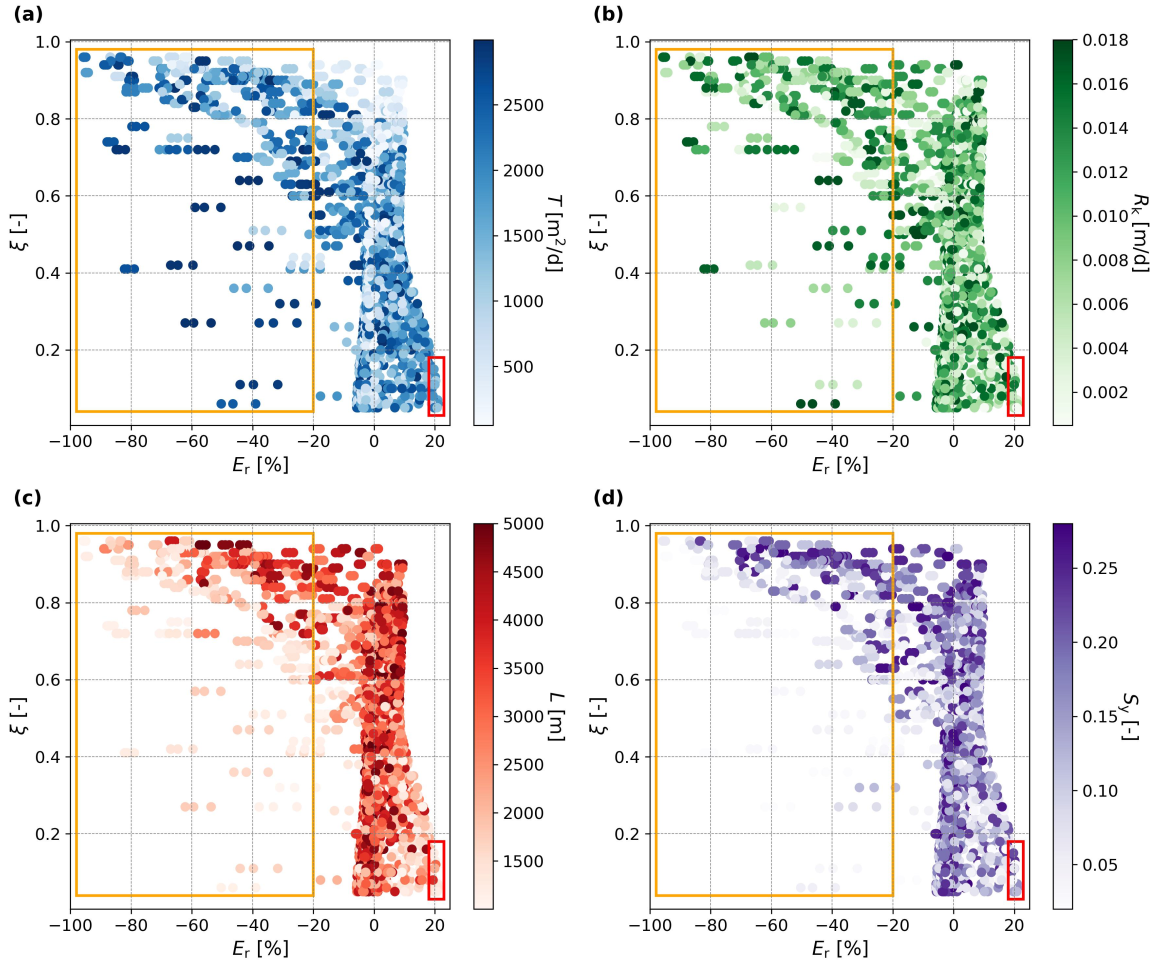


Figure A1: Scatter plots illustrating relationships between E_r and ξ for the linear LRC approach: (a) E_r vs ξ with respect to T, (b) E_r vs ξ with respect to R_k, (c) E_r vs ξ with respect to L, (d) E_r vs ξ with respect to S_y. The orange and red rectangles highlight large error values (represented as individual points in Figure 5) that underestimate and overestimate recharge, respectively.

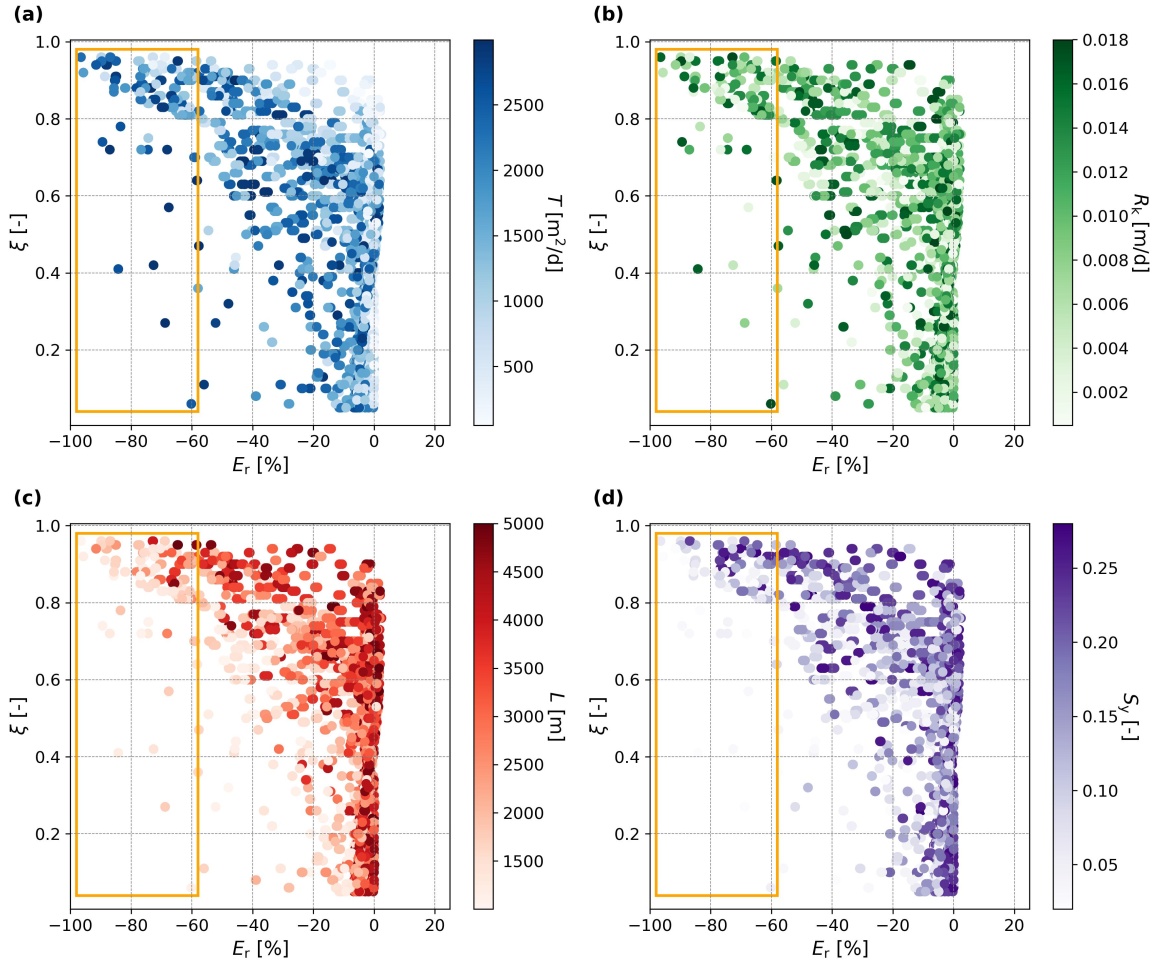


Figure A2: Scatter plots illustrating relationships between E_r and ξ for the power LRC approach: (a) E_r vs ξ with respect to T, (b) E_r vs ξ with respect to R_k, (c) E_r vs ξ with respect to L, (d) E_r vs ξ with respect to S_y. The orange rectangle highlights large error values (represented as individual points in Figure 5) that underestimate recharge.

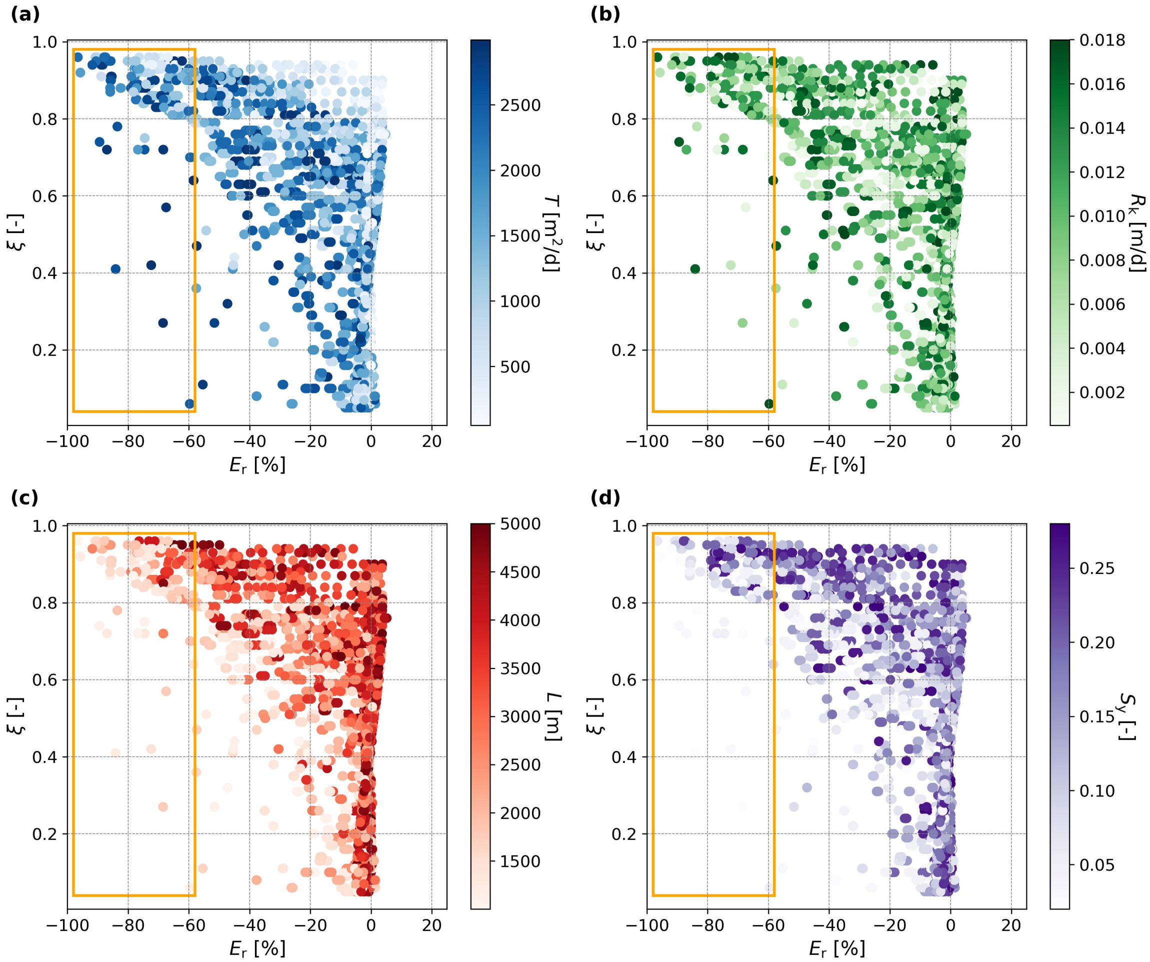


Figure A3: Scatter plots illustrating relationships between E_{r} and ξ for the event-based MRC approach: (a) E_{r} vs ξ with respect to T, (b) E_{r} vs ξ with respect to R_{k}, (c) E_{r} vs ξ with respect to L, (d) E_{r} vs ξ with respect to S_{y}. The orange rectangle highlights large error values (represented as individual points in Figure 5) that underestimate recharge.

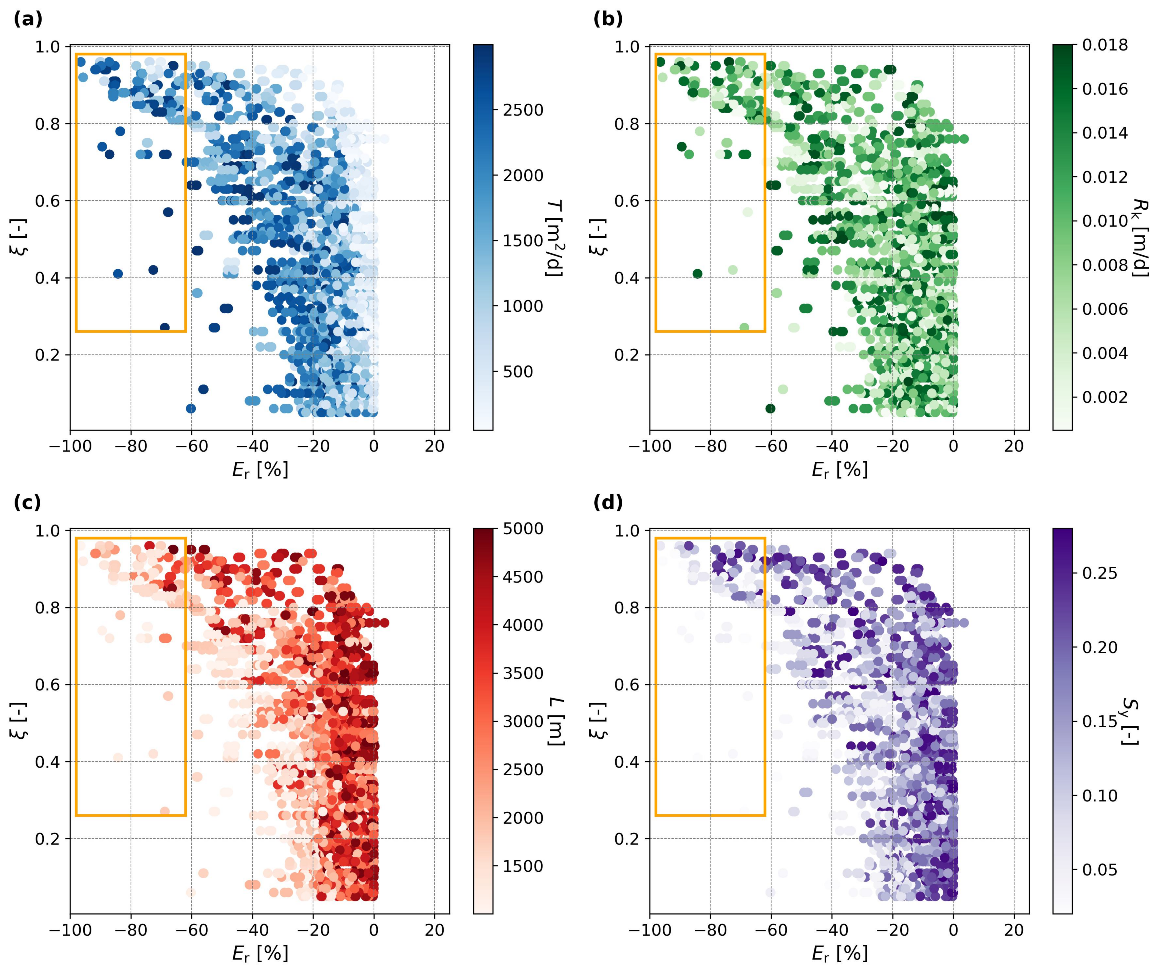
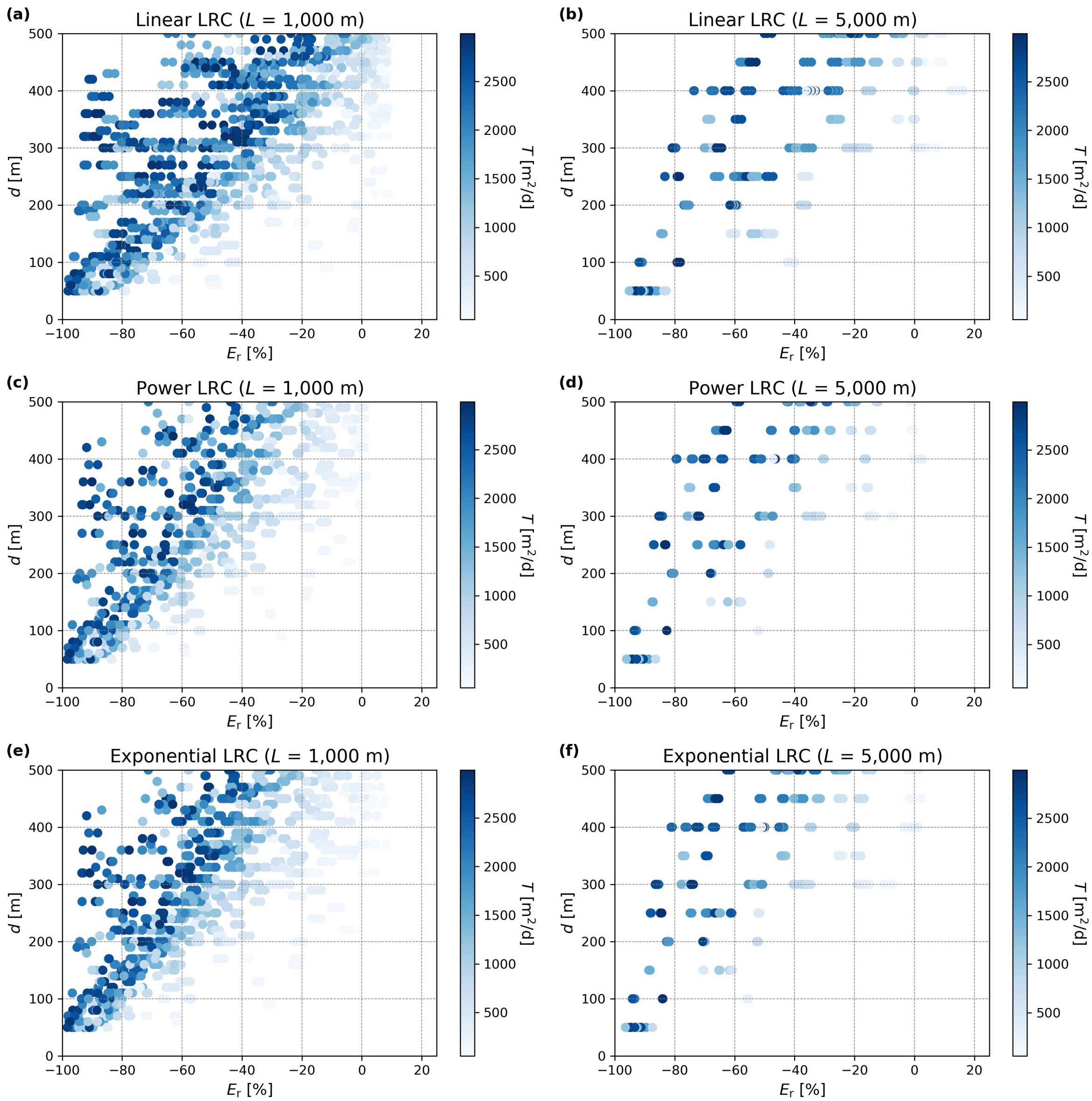


Figure A4: Scatter plots illustrating relationships between E_{r} and ξ for the RISE method: (a) E_{r} vs ξ with respect to T, (b) E_{r} vs ξ with respect to R_{k}, (c) E_{r} vs ξ with respect to L, (d) E_{r} vs ξ with respect to S_{y}. The orange rectangle highlights large error values (represented as individual points in Figure 5) that underestimate recharge.


(a)
Linear LRC (L = 1,000 m)
(b)
Linear LRC (L = 5,000 m)
(c)
Power LRC (L = 1,000 m)
(d)
Power LRC (L = 5,000 m)
(e)
Exponential LRC (L = 1,000 m)
(f)
Exponential LRC (L = 5,000 m)
d [m]
E_r [%]
T [m^2/d]
500
400
300
200
100
0
−100
−80
−60
−40
−20
0
20
2500
2000
1500
1000
500

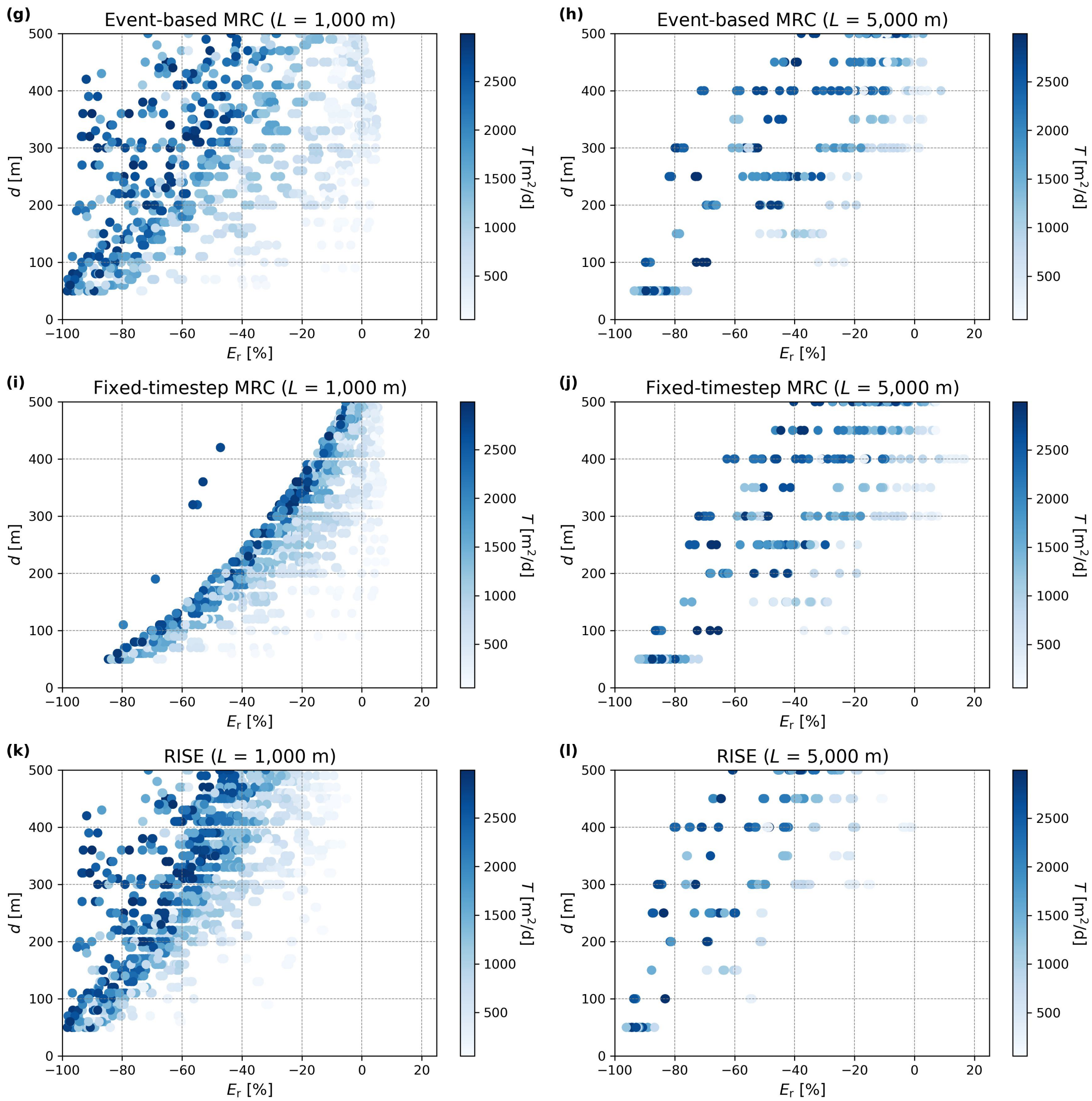


Figure A5: Scatter plots illustrating relationships between E_r, d, and T for all WTFM variants where L was fixed to 1,000 m and 5,000 m.

References

[1] J. F. Águila, J. Samper, and B. Pisani. Parametric and numerical analysis of the estimation of groundwater recharge from water-table fluctuations in heterogeneous unconfined aquifers. *Hydrogeology Journal*, 27(4):1309–1328, 2019. doi: 10.1007/s10040-018-1908-x.

[2] S. Barua, I. Cartwright, P. E. Dresel, and E. Daly. Using multiple methods to investigate the effects of land-use changes on groundwater recharge in a semi-arid area. *Hydrology and Earth System Sciences*, 25(1):89–104, 2021. doi: 10.5194/hess-25-89-2021.

[3] A. L. Becke, S. C. Solórzano-Rivas, and A. D. Werner. The watertable fluctuation method of recharge estimation: A review. *Advances in Water Resources*, 189:104635, 2024. doi: 10.1016/j.advwatres.2024.104635.

[4] I. Cartwright, U. Morgenstern, H. Hofmann, and B. Gilfedder. Comparisons and uncertainties of recharge estimates in a temperate alpine catchment. *Journal of Hydrology*, 590:125558, 2020. doi: 10.1016/j.jhydrol.2020.125558.

[5] R. S. Crosbie, P. Binning, and J. D. Kalma. A time series approach to inferring groundwater recharge using the water table fluctuation method. *Water Resources Research*, 41 (1):2004WR003077, 2005. doi: 10.1029/2004WR003077.

[6] R. S. Crosbie, I. D. Jolly, F. W. Leaney, and C. Petheram. Can the dataset of field based recharge estimates in Australia be used to predict recharge in data-poor areas? *Hydrology and Earth System Sciences*, 14(10):2023–2038, 2010. doi: 10.5194/hess-14-2023-2010.

[7] R. S. Crosbie, R. C. Doble, C. Turnadge, and A. R. Taylor. Constraining the Magnitude and Uncertainty of Specific Yield for Use in the Water Table Fluctuation Method of Estimating Recharge. *Water Resources Research*, 55(8):7343–7361, 2019. doi: 10.1029/2019WR025285.

[8] M. O. Cuthbert. Straight thinking about groundwater recession. *Water Resources Research*, 50(3):2407–2424, 2014. doi: 10.1002/2013WR014060.

[9] G. N. Delin, R. W. Healy, D. L. Lorenz, and J. R. Nimmo. Comparison of local- to regional-scale estimates of ground-water recharge in Minnesota, USA. *Journal of Hydrology*, 334(1-2):231–249, 2007. doi: 10.1016/j.jhydrol.2006.10.010.

[10] C. Gong, P. G. Cook, R. Therrien, W. Wang, and P. Brunner. On Groundwater Recharge in Variably Saturated Subsurface Flow Models. *Water Resources Research*, 59(9):e2023WR034920, 2023. doi: 10.1029/2023WR034920.

[11] A. Gumuła-Kawecka, B. Jaworska-Szulc, A. Szymkiewicz, W. Gorczewska-Langner, M. Pruszkowska-Caceres, R. Angulo-Jaramillo, and J. Šimůnek. Estimation of groundwater recharge in a shallow sandy aquifer using unsaturated zone modeling and water table fluctuation method. *Journal of Hydrology*, 605:127283, 2022. doi: 10.1016/j.jhydrol.2021.127283.

[12] B. Hagedorn, A. I. El-Kadi, A. Mair, R. B. Whittier, and K. Ha. Estimating recharge in fractured aquifers of a temperate humid to semiarid volcanic island (Jeju, Korea) from water table fluctuations, and Cl, CFC-12 and 3H chemistry. *Journal of Hydrology*, 409 (3-4):650–662, 2011. doi: 10.1016/j.jhydrol.2011.08.060.

[13] C. R. Harris, K. J. Millman, S. J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020. doi: 10.1038/s41586-020-2649-2.

[14] R. W. Healy and P. G. Cook. Using groundwater levels to estimate recharge. *Hydrogeology Journal*, 10(1):91–109, 2002. doi: 10.1007/s10040-001-0178-0.

[15] R. W. Healy and B. R. Scanlon. *Estimating Groundwater Recharge*. Cambridge University Press, Cambridge, 2010. URL https://doi.org/10.1017/CBO9780511780745.

[16] C. S. Heppner and J. R. Nimmo. A computer program for predicting recharge with a master recession curve. Scientific Investigations Report 2005-5172, U.S. Geological Survey, California, 2005. URL https://doi.org/10.3133/sir20055172.

[17] V. Hung Vu and B. J. Merkel. Estimating groundwater recharge for Hanoi, Vietnam. *Science of The Total Environment*, 651:1047–1057, 2019. doi: 10.1016/j.scitotenv.2018.09.225.

[18] Z. Jie, J. van Heyden, D. Bendel, and R. Barthel. Combination of soil-water balance models and water-table fluctuation methods for evaluation and improvement of groundwater recharge calculations. *Hydrogeology Journal*, 19(8):1487–1502, 2011. doi: 10.1007/s10040-011-0772-8.

[19] A. I. Johnson. Specific yield: compilation of specific yields for various materials. U.S. Geological Survey Water Supply Paper 1662-D, U.S. Government Printing Office, Washington D.C., 1967. URL https://pubs.usgs.gov/publication/wsp1662D.

[20] A. C. King, M. Raiber, M. E. Cox, and D. I. Cendón. Comparison of groundwater recharge estimation techniques in an alluvial aquifer system with an intermittent/ephemeral stream (Queensland, Australia). *Hydrogeology Journal*, 25(6):1759–1777, 2017. doi: 10.1007/s10040-017-1565-5.

[21] S. Lanini. ESPERE Version 2 User's Guide. U.S. Geological Survey Water Supply Paper, French Geological Survey, France, 2020. URL https://pubs.usgs.gov/publication/wsp1472.

[22] M. Maasland. Water table fluctuations induced by intermittent recharge. *Journal of Geophysical Research*, 64(5):549–559, 1959. doi: 10.1029/JZ064i005p00549.

[23] J. R. Nimmo and K. S. Perkins. Episodic Master Recession Evaluation of Groundwater and Streamflow Hydrographs for Water-Resource Estimation. *Vadose Zone Journal*, 17 (1):1–25, 2018. doi: 10.2136/vzj2018.03.0050.

[24] J. R. Nimmo, C. Horowitz, and L. Mitchell. Discrete-Storm Water-Table Fluctuation Method to Estimate Episodic Recharge. *Groundwater*, 53(2):282–292, 2015. doi: 10.1111/gwat.12177.

[25] W. C. Rasmussen and G. E. Andreasen. Hydrologic budget of the Beaverdam Creek basin, Maryland. U.S. Geological Survey Water Supply Paper 1472, United States Geological Survey, Washington, 1959. URL https://pubs.usgs.gov/publication/wsp1472.

[26] B. R. Scanlon, R. W. Healy, and P. G. Cook. Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, 10:18 – 39, 2002. doi: 1007/s10040-0010176-2.

[27] S. C. Solórzano-Rivas, A. D. Werner, and N. I. Robinson. Assessing the reliability of exponential recession in the water table fluctuation method. *Advances in Water Resources*, 193:104821, 2024. doi: 10.1016/j.advwatres.2024.104821.

[28] N. Somaratne, K. Smettem, and J. Frizenschaf. Three criteria reliability analyses for groundwater recharge estimations. *Environmental Earth Sciences*, 72(6):2141–2151, 2014. doi: 10.1007/s12665-014-3121-0.

[29] J. Tóth. A theoretical analysis of groundwater flow in small drainage basins. *Journal of Geophysical Research*, 68(16):4795–4812, 1963. doi: 10.1029/JZ068i016p04795.

[30] S. Viaroli, D. Di Curzio, D. Lepore, and R. Mazza. Multiparameter daily time-series analysis to groundwater recharge assessment in a caldera aquifer: Roccamonfina Volcano, Italy. *Science of The Total Environment*, 676:501–513, 2019. doi: 10.1016/j.scitotenv.2019.04.327.

[31] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G.-L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. De Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, and Y. Vázquez-Baeza. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020.

[32] J. von Freyberg, C. Moeck, and M. Schirmer. Estimation of groundwater recharge and drought severity with varying model complexity. *Journal of Hydrology*, 527:844–857, 2015. doi: 10.1016/j.jhydrol.2015.05.025.

[33] X. Wang, G. Zhang, and Y. J. Xu. Spatiotemporal groundwater recharge estimation for the largest rice production region in Sanjiang Plain, Northeast China. *Journal of Water Supply: Research and Technology-Aqua*, 63(8):630–641, 2014. doi: 10.2166/aqua.2014.024.

[34] E. Wendland, C. Barreto, and L. H. Gomes. Water balance in the Guarani Aquifer outcrop zone based on hydrogeologic monitoring. *Journal of Hydrology*, 342(3-4):261–269, 2007. doi: 10.1016/j.jhydrol.2007.05.033.

[35] W. Wilopo and D. P. E. Putra. Groundwater recharge estimation using groundwater level fluctuation patterns in unconfined aquifer of Yogyakarta City, Indonesia. *Kuwait Journal of Science*, 48(2), 2021. doi: 10.48129/kjs.v48i2.9397.

[36] A. Yenehun, F. Nigate, A. S. Belay, M. T. Desta, M. Van Camp, and K. Walraevens. Groundwater recharge and water table response to changing conditions for aquifers at different physiography: The case of a semi-humid river catchment, northwestern highlands of Ethiopia. *Science of The Total Environment*, 748:142243, 2020. doi: 10.1016/j.scitotenv.2020.142243.

[37] Y.-K. Zhang and K. E. Schilling. Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: A Field observation and analysis. *Journal of Hydrology*, 319(1-4):328–338, 2006. doi: 10.1016/j.jhydrol.2005.06.044.