Received: 24 February 2018 | Revised: 21 August 2018 | Accepted: 22 August 2018
DOI: 10.1002/rra.3357

RESEARCH ARTICLE

WILEY

A predictive management tool for blackfly outbreaks on the Orange River, South Africa

Nicholas A. Rivers-Moore[1] | Trevor R. Hill[2]

[1] Centre for Water Resources Research, University of KwaZulu-Natal, Scottsville, South Africa

[2] Discipline of Geography, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Scottsville, South Africa

Correspondence
Nicholas A. Rivers-Moore, Centre for Water Resources Research, University of KwaZulu-Natal, Private Bag X01, Scottsville, 3209, South Africa.
Email: blackfly1@vodamail.co.za

Funding information
Red Meat Research & Development South Africa; Water Research Commission, Grant/Award Number: K5/2459

Abstract

Downstream flow alteration resulting from river impoundment or interbasin transfer schemes, while improving water supply assurance levels, has been shown to have negative ecological consequences, including outbreaks of pest blackfly. In South Africa's Orange River, large impoundments constructed in the 1970s have created an ongoing blackfly outbreak problem. Although the severity of the outbreaks has been successfully managed using aerial applications of larvicides, periodic outbreaks continue to occur. Understanding the interactions of the multiple variables driving the outbreaks is complex. We integrated variables useful in predicting outbreak conditions (discharge, water temperature, seston concentration, benthic algae) using a Bayesian network approach. Data to define probabilities were collected at 11 sites over four sampling seasons, and system states were derived using flow and water temperature thresholds. The late summer months (February, March, and April) were most favourable for pest blackly outbreaks, and the probability of an outbreak is six times higher for postimpoundment versus preimpoundment flow conditions. The model was successful in integrating multiple environmental variables that act as triggers for pest blackfly outbreaks. The efficacy of the model as a management tool will increase if ongoing monitoring data are incorporated into the model as case files.

KEYWORDS

Bayesian network, flow, probability, *Simulium chutteri*, turbidity

1 | INTRODUCTION

Filter-feeding blackfly larvae are ubiquitous in lotic systems and inhabit a broad spectrum of hydraulic habitats from small clear mountain streams to large turbid rivers (Crosskey, 1973). Notwithstanding the economic benefits of flow assurance, impoundment of large river systems may result in overwhelming dominance of aquatic macroinvertebrate communities by pest species such as blackfly (O'Keeffe & de Moor, 1988). This is a consequence of downstream changes in natural flow regimes, through either homogenization and reversal of hydrographs, or flow augmentation in receiving systems from interbasin transfer schemes (Rivers-Moore, de Moor, Morris, & O'Keeffe, 2007; Snaddon & Davies, 1998). Such ecological consequences are not always anticipated or internalized in cost–benefit analyses supporting the development of such schemes. Examples of this include large rivers in West Africa, the Saskatchewan River system in Canada, and the Orange and Great Fish Rivers of South Africa (Chalifour, Boisvert, & Back, 1990; Fredeen, 1977; Rivers-Moore et al., 2007; Rivers-Moore, Palmer, & Dallas, 2014), where various permutations of negative influences on human health, livestock and poultry farming, agriculture, and recreation have manifested.

In South Africa, blackfly outbreaks along the middle and lower Orange River have the potential to cause losses to livestock production estimated at US$13.3 million per annum in 2013 (Rivers-Moore et al., 2014). This figure is a conservative estimate as it excludes losses in the tourism and irrigated agricultural sectors through lost revenue and labour days (Mullins, 2007). Economic losses occur along approximately 1,200 km of the middle and lower reaches of the Orange River (Palmer, 1997). This is the river segment downstream of Van Der Kloof Dam, a major impoundment regulating flows in the Orange River. The major pest species is *Simulium chutteri*; however, *Simulium damnosum*, *Simulium nigritarse*, and *Simulium adersi* are also of concern (de Moor, 1994, and citing others). Switching between clear and turbid conditions, which are driven by changes in flow volumes and seston concentration, favour either the major pest species (*S. chutteri* and *S. damnosum*) under more turbid conditions or less problematic species (*S. adersi* and *S. nigritarse*) under clearer conditions (Rivers-Moore & Palmer, 2018). Whereas the latter feed predominantly on birds, the former major pest species feed primarily on mammals and are therefore considered a management concern due to their negative impact to sheep farmers and their nuisance value to agricultural labourers (Rivers-Moore & Palmer, 2018).

There are a number of options for blackfly control, including flow manipulation, physical removal of aquatic weeds, and aerial spraying of adult flies; protection of livestock using insecticides; and biological control and larvicide application. The Blackfly Control Programme along the middle and lower Orange River was started in 1991, based on aerial applications of larvicides to control the key pest species *S. chutteri*. Effective larvicide concentrations have been established through research, with larvicide volumes calculated at each application based on measured discharge. The volumes of larvicide used, together with relatively large discharges (mean = 100 $m^3 \cdot s^{-1}$), mean that downstream carry of active ingredients from upstream applications results in an overapplication of larvicides (Chalifour et al., 1990; Rivers-Moore, Bangay, & Palmer, 2008). It extends over some 850 km of the middle and lower Orange River, where 148 rapids have been identified as optimal breeding habitat for pest blackfly species (Palmer, Rivers-Moore, Mullins, McPherson, & Hattingh, 2007). Larvicides are usually applied three times in autumn and six times in spring (Palmer & Palmer, 1995). Larvicides registered for blackfly control in South Africa are Teknar® and VectoBac® (produced from the naturally occurring bacteria *Bacillus thuringiensis* var. *israelensis*) and Abate® (organophosphate temephos; Palmer & Palmer, 1995). The former bacterial larvicides are high target specific to simuliid Diptera, whereas the organophosphates are less selective (Palmer & Rivers-Moore, 2008). Wide-scale application of Abate has led to resistance being developed in the major pest blackfly species Orange River populations as a result of prolonged exposure to this larvicide (Palmer & Palmer, 1995; Palmer & Rivers-Moore, 2008). Consequently, the current control programme is restricted to, and completely dependent on, correct application of *B. thuringiensis* var. *israelensis*. The success of the control programme depends largely on the correct timing of larvicide applications, informed through a monitoring programme using a 10-point scoring system for larval and pupal densities developed by Palmer (1994). Larval density data are scored fortnightly by staff from the regional South African Department of Agriculture, Forestry and Fisheries.

There has been a research history of more than 30 years in response to the "blackfly problem" along the Orange River. Projects have included fundamental research of blackfly ecology on the Orange River, to inform the design of the Blackfly Control Programme by Palmer (1997), and a follow-up project 10 years later to explore alternative larvicides due to larval resistance of temephos (Palmer et al., 2007). However, despite a long history of research, monitoring, and management, periodic outbreaks of blackfly continue to occur, with the most recent outbreak in 2011 (Rivers-Moore et al., 2014), and before that in 2000–2001 (Palmer et al., 2007). Reasons for periodic outbreaks include sporadic higher than normal winter flows (Palmer et al., 2007), changes in turbidity levels promoting switching of dominant blackfly species (Fredeen, 1977; Rivers-Moore et al., 2014), and larvicidal resistance (Palmer & Rivers-Moore, 2008).

Integrating the interacting effects of all of these variables is a challenge. Bonkewitzz and Palmer (1997) developed an interactive, flexible, rule-based probabilistic model for river managers involved with the Orange River Blackfly Control Programme. Understanding the causal factors contributing to blackfly outbreaks requires knowledge of the variables governing blackfly larval numbers, which include water velocity, available habitat, and water temperature (Crosskey, 1973; de Moor, 2003; Palmer & O'Keeffe, 1995; Rivers-Moore, de Moor, Birkholz, & Palmer, 2006; Rivers-Moore, Hughes, & de Moor, 2008; Sheldon & Oswood, 1977). Given the interacting dynamics of each variable, and the uncertainty around the relative contributions of each variable under different seasons, a suitable framework for representing the problem is a Bayesian network (BN). This is a probabilistic network for reasoning under uncertainty, wherein current knowledge on multiple variables and their interdependencies are quantified and graphically represented into a probabilistic modelling framework (Jensen & Nielsen, 2007; Kjaerulff & Madsen, 2008). The probability of an event is conditional on other factors (Jensen & Nielsen, 2007), where "each fact is suggestive in itself [and] together they have a cumulative force" (Sherlock Holmes: "The Adventure of the Bruce-Partington Plans"). BNs are causal networks with the strength of links represented as conditional probabilities and useful in calculating new probabilities as new information becomes available.

Rivers-Moore et al. (2014) developed a simple BN model for two species on the Orange River, although this model did not incorporate the effects of water temperature or turbidity on outbreak probabilities. Such networks are useful as a decision support tool for considering the influences of multiple variables on a measured response variable (Stewart-Koster et al., 2010; for example, blackfly outbreaks). A BN essentially consists of cause-and-effect relationships and is a tool for facilitating the development of conceptual models for representing relationships among variables, even if the relationships involve uncertainty. Consequently, this approach is free from the arguments of too little data, and BNs show strong prediction accuracy even using small sample sizes (Batchelor & Cain, 1999; Kjaerulff & Madsen, 2008; Uusitalo, 2007). With the use of this approach, BNs are particularly useful in predicting the likelihood of an event occurring through consideration of both independent and interactive (conditional) causal environmental variables on the response variable —in this case, probability of blackfly outbreak (Stewart-Koster et al., 2010). This paper develops a probabilistic model for predicting when outbreaks are likely to be most severe, and which environmental variables are most likely to be responsible for them. Together, these provide a tool for assisting with the seasonal planning efforts of the Blackfly Control Programme and informing which variables should be monitored.

2 | METHODS

2.1 | Study sites

Eleven sites along the middle and lower Orange River, distributed along some 600 km downstream of Van Der Kloof Dam, were selected for collection of primary data (Figure 1; Table 1). Sites extended over an elevation range of 600 m and were chosen to represent both single- and multiple-channelled river reaches. Channel type was previously identified by Rivers-Moore et al. (2014) as affecting flow rates and current velocities. Multiple channel sections were identified using GoogleEarth™ and plotted on the river profile. The spatial distribution of sections with multiple channels was characterized on the basis of the association between points (degree of clustering vs. regular spacing). The co-ordinates of all instream barriers and downstream distances between point pairs were calculated for $2 \times n$ and $n \times n$ matrices. These matrices were used in second-order analyses, a suitable technique for assessing clustering of points in one or more dimensions (Fortin & Dale, 2005; Rosenberg & Anderson, 2011). Outputs are modified Ripley's K values that are a function of how many points fall within a series of different radius values for each point.

2.2 | Data collection

Sampling was undertaken in late spring (November 2015), late summer (March 2016), winter (July 2016), and early summer (December 2016). Moving from downstream to upstream so as not to contaminate or trample downstream sites, we sampled across a range of hydraulic habitats and reeds where immature blackfly (larvae and pupae) were expected. Although there is evidence to support limited drift of larvae downstream within a breeding site (Rivers-Moore et al., 2007), it is highly unlikely that cross-contamination between sampling sites occurred given the sessile nature of blackfly larvae, and the average interriffle distances of 5.7 km (Rivers-Moore, Bangay, & Palmer, 2008). Samples were collected either from reeds cut and preserved or using a 250-μm mesh net downstream of fist-sized rocks. Larval densities were rated according to the 10-point scale of Palmer (1994). Blackfly pupae and larvae were collected and preserved in 70% ethanol. Each sample was identified to species level, and relative abundances were recorded in the laboratory using the taxonomic keys of de Moor (2003). All data for each sample at each site and per seasonal sampling event were collated into spreadsheet data matrices, with associated hydraulic and water quality data and manipulated using pivot tables. Raw abundance data per species and life history stage, and presence and absence data were used in analyses. Blackfly species turnover between seasons and sites was compared using a Bray–Curtis analysis (McCune & Mefford, 2011). Reed areas extending 100 m upstream and downstream at all sites were calculated using on-screen digitizing of satellite images from GoogleEarth™.

Turbidity (cm) was measured using a clarity tube, in association with the presence/absence of algae, which reduces blackfly habitat. Turbidity values were converted to seston concentrations (mg·L^{-1}; Equation (1); Palmer, 1997; Rivers-Moore et al., 2007). Spot readings of pH and conductivity (μS·cm^{-1}) were recorded using a Hanna pH/conductivity meter.

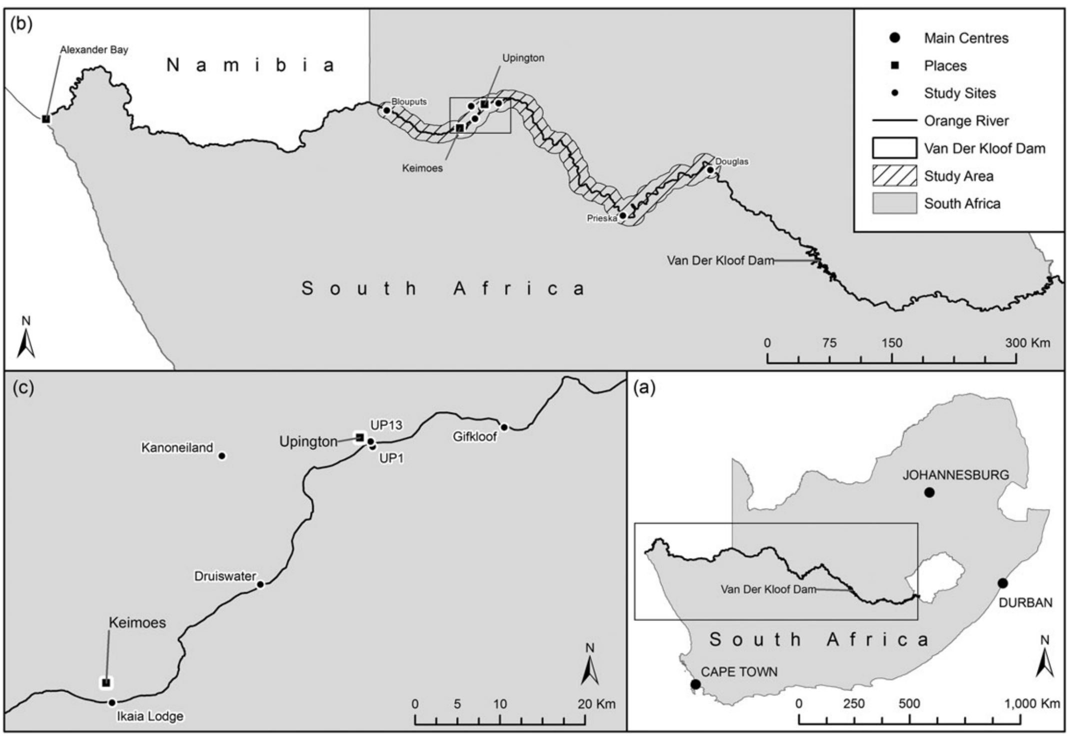

FIGURE 1 Map of study area, showing full extent of study area (a) and location of study sites relative to Van Der Kloof Dam (b). The blackfly control zone extends from Douglas downstream to some 400 km below Blouputs. Higher resolution of sites between Gifkloof and Keimos is show in (c)

TABLE 1 Study site details, including downstream distance from Van Der Kloof Dam (DD), and elevation

Site	Site name	Latitude °S	Longitude °E	DD (km)	Elevation (m amsl)	Channel type
1	Douglas	−29.16194	23.69623	174.6	993	Single
2	Prieska	−29.65553	22.74592	355.1	926	Single
3	Gifkloof	−28.43743	21.40092	621.6	800	Single
4	UP1	−28.45798	21.26165	636.5	785	Single
5	UP13	−28.45262	21.25943	636.5	785	Single
6	Druiswater	−28.60385	21.14277	660.3	778	Single
7	Kanoneiland	−28.46768	21.10197	666.5	765	Single
8	UP8	−28.68780	21.06878	671.8	746	Single
9	UP12	−28.69490	21.01452	679.3	726	Multiple
10	Ikaia Lodge	−28.72913	20.98595	683.8	724	Single
11	Blouputs	−28.51377	20.18694	785.0	439	Single

$$TSS = \exp\left(\frac{\log\frac{SD}{256}}{-0.616}\right), \quad (1)$$

where *TSS* is total suspended solids ($mg \cdot L^{-1}$) and *SD* is clarity (cm).

Water temperatures are a controlling variable for larvicidal efficacy (Palmer, 1997), algal mats (de Moor, 1994), and blackfly larval development (de Moor, 1982, 1994). Hourly water temperature data were collected at all sites using Hobo TidbiT v2 water temperature loggers and aligned for a common period from November 4, 2015, to November 2, 2016. Data were processed into metrics describing magnitudes, frequencies, durations, and timing of thermal events (Rivers-Moore, Dallas, & Morris, 2013), based on thermal thresholds for *S. chutteri* of 10°C, 26.67°C, and 30°C, which describe a lower threshold below which pupation does not occur, a chronic stress threshold, and an upper thermal threshold (Rivers-Moore, Dallas, & Ross-Gillespie, 2013). Definition of thermal seasons of mean daily water temperatures was determined using regime shift detection software (Rodionov, 2006; $p < 0.01$; cut-off length = 30; Huber's weight parameter = 1).

For the hydraulic data, current velocity was measured at each sampling point, using a transparent velocity head rod. Differentials in depth between the current "head" and the lower depth were converted to velocities. Sampling point depths (cm) were recorded using a depth stick. For the hydrological data, observed mean daily flow data time series were obtained from the national Department of Water and Sanitation's Hydrological Information System (www.dwaf.gov.za/Hydrology). We used flow gauging data from two stations with the longest time series data available: D3H008/Marksdrift: 1935–2016; and D7H008/Upington: 1942–2016. A critical discharge threshold of 100 $m^3 \cdot s^{-1}$ was derived from the velocity–discharge relationships in Palmer (1997) and using a critical velocity of 1 $m \cdot s^{-1}$ for both *S. chutteri* and *S. damnosum* (Palmer & Craig, 2000; Rivers-Moore et al., 2007). Return intervals of flows exceeding 100 $m^3 \cdot s^{-1}$ were calculated for preimpoundment and postimpoundment periods (1942–1977; 1978–2016).

2.3 | Bayesian network model

Data from the 11 study sites were used as the basis for developing the BN model. For the sake of model parsimony, the BN model objective node was restricted to two node states, namely, the "major" problem species (*S. chutteri* and *S. damnosum*) and the "minor" problem species (*S. adersi* and *S. nigritarse*).

BN models are not explicitly able to reflect temporal or spatial patterns (Cain, 2001). The approach to dealing with this was to build as many models as required to represent study area spatial units, and relevant time periods. Sites were grouped according to seasonal values of pH, turbidity (cm), and conductivity ($\mu S \cdot cm^{-1}$). These data were analysed using a principal component analysis (correlation matrix), and sites classified using a cluster analysis (Euclidean distance measure; group averaging linkage method). In the second approach, sites were grouped according to thermal metrics describing water temperature time series, on the basis of the methods of Rivers-Moore, Dallas, and Morris (2013), and the thermal data collected in this study.

The BN model was designed to take the following into account:

- benthic algae dominating the substrate habitat and controlled by water temperature and seston concentration (de Moor, 1994);
- larvicidal efficacy as a function of water temperature and seston concentration (Palmer, 1997);
- outbreak probability as a function of dominance of the "strong porous" species and affected by water temperature; and
- dominance of the pest blackfly species complex determined by the moderating nodes of "abiotic" and "biotic" conditions dominating (de Moor, 1994).

Nodes were linked in cause-and-effect sequences using the Bayesian software Netica v 4.16 (Norsys Software Corporation, 2010) and assigned variable states. All variables used discrete states, arranged from most positive to most negative. Development of a BN was an iterative process of testing the logic of relationships and keeping the network as parsimonious as possible. No more than three parent nodes (a variable with links going out to other variables; Cain, 2001) were linked to any child node (a variable with links to it from other variables; Cain, 2001), as the elements of a conditional probability table increase exponentially according to i^n based on number of states (i) and the number of parent nodes (n; Cain, 2001). The water temperature threshold was linked to a number of child nodes each likely to have a particular temperature threshold; a trade-off was made in assigning a 20°C mean daily water temperature threshold as being a compromise threshold to serve all three variables. With the use of data from the field surveys, case files were generated, with each data point representing a data record. Simple logic statements are used for five of the nodes (biotic, abiotic, larvicidal efficacy, pest complex, and outbreak probability), based on combinations of their parent node states. For example, "Discharge = High AND Seston Concentration = High" resulted in the abiotic node being assigned a "favourable" state. Node states were defined according to the threshold values, and conditional probabilities were generated from the case files.

To validate the model, two approaches were used. In the first, we divided the case file data into training and test case files, on a 3:1 ratio. Data records were assigned random numbers sorted from smallest to largest and were selected for each exercise. The test case file was applied to both the upper and lower Orange River BN models. In the second approach, the probabilities of each state of the parent nodes were defined on the basis of a number of approaches; for example, for flow conditions, the return intervals for flows below or exceeding 100 $m^3{\cdot}s^{-1}$ were calculated from flow data. This threshold corresponds with optimal velocity thresholds for both *S. chutteri* and *S. damnosum* (Palmer, 1997; Palmer & Craig, 2000). Return intervals were calculated for two periods defined according to when the major impoundment controlling downstream flows on the Orange River was completed, namely, 1942–1977 versus 1978–2016, and probabilities calculated for each state. An exponential relationship between flow rates and seston concentration was used to calculate seston concentration time series for preimpoundment and postimpoundment flow conditions, from which return intervals were calculated for a threshold value of 60 $mg{\cdot}L^{-1}$ (Palmer, 1997; Equation (2)). Monthly outbreak probabilities were compared against historical data of adult blackfly annoyance levels. The fly worry index was previously used in the 1990s (Palmer, 1997) as a 4-point scoring system reflecting the annoyance levels of adult blackfly. It is assumed that there will be a 1- to 2-week time lag between the larval density data and the adult fly worry index data, on the basis of the temperature-dependent time lag between larval and adult life history stages.

$$Seston = 1.92{*}Flow^{0.755}. \quad (2)$$

3 | RESULTS

A total of 25 multiple-channelled sections were identified on the middle and lower Orange River. Multiple channel zones were strongly clustered within 60- to 80-km segments, with clusters regularly spaced at larger scales (Figure 2).

All water quality data exhibited seasonal variation (Figure 3), with little downstream gradient difference. Water clarity was generally low (<42 cm) and the highest during the July 2016 survey, with correspondingly high levels of algae on the rocks. Based on the seston concentration versus flow volume curves, low seston concentrations occurred at low flows, with distinct seasonal patterns. Values of pH reflected neutral to slightly alkaline conditions at all sites (7.0–8.5). Conductivity values ranged from 400 to 600 $\mu S{\cdot}cm^{-1}$, being highest

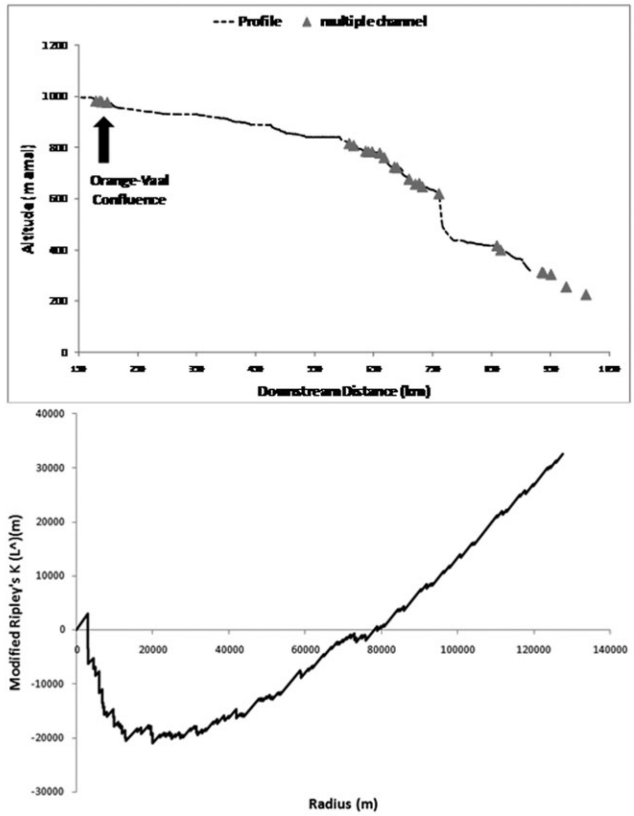

FIGURE 2 Orange River profile based on downstream distance from Van Der Kloof Dam showing distribution of multiple-channelled sites on the Orange River downstream of Van Der Kloof Dam (top); results of Ripley's *K* analysis on one-dimensional data of barriers along the lower and middle Orange River. Values of zero reflect random distributions, values < 0 indicate clumping, and values > 0 indicate regular spacing (bottom)

during the July 2016 survey, corresponding with relatively low flow volumes.

Water temperatures exhibited a marked cooling trend from late March 2016, with Sites 1–2 slightly cooler than the remaining downstream sites (3,600 vs. 4,000 degree days > 10°C per annum). This was predominantly a consequence of a marked reduction in daily water temperature ranges and a cooling trend from late March. The annual thermograph showed 10 significant changes in mean daily water temperatures (Figure 4).

Seasonal plots of the relative proportions of contribution to overall sample numbers demonstrate clear switching of dominant species between sites and seasons, but with specific site clusters based on species abundances (Figure 5). Maximum abundances relative to velocity values showed clear species-specific responses, with *S. chutteri* and *S. damnosum* abundances peaking at 1.2 $m{\cdot}s^{-1}$. In contrast, *S. adersi* and *S. nigritarse* showed preferences for velocities of 0.6 and 0.8 $m{\cdot}s^{-1}$, respectively (Figure 6; Table 2). There was little relationship between blackfly species abundances and pH or conductivity.

3.1 | Bayesian network model

In terms of water quality, all sites exhibited similar characteristics, with little evidence to support grouping of sites based on water quality.

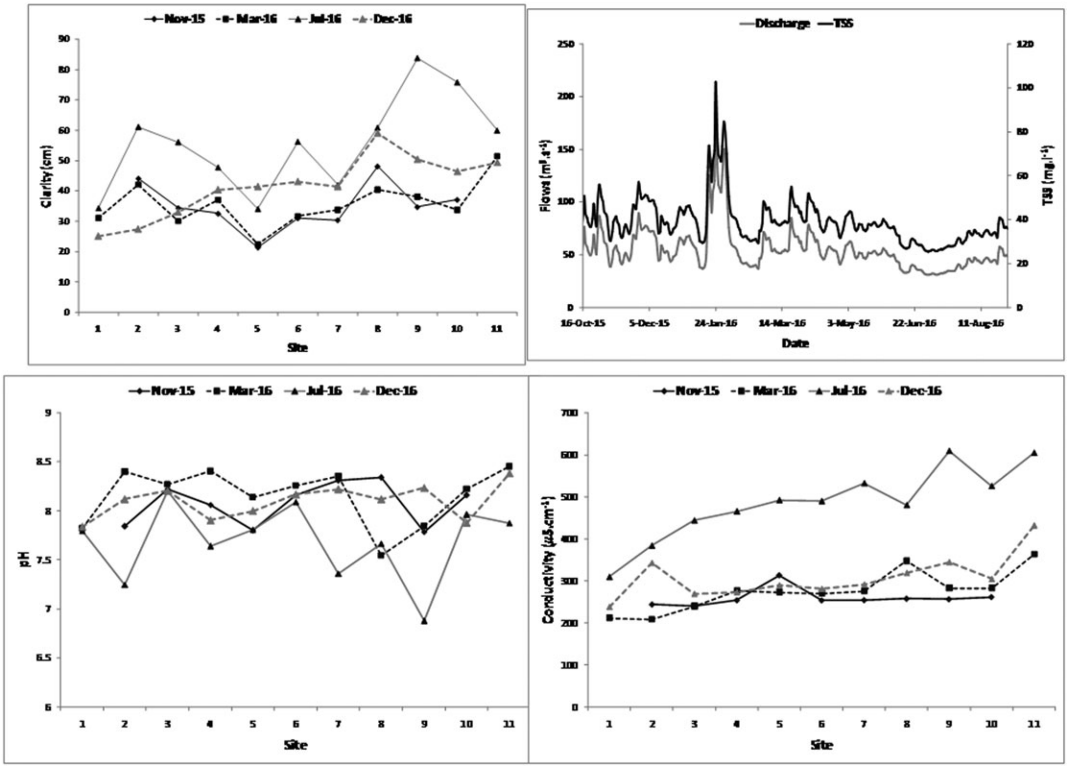

FIGURE 3 Clarity (top, left); pH (bottom, left); and conductivity values for late spring (November 2015), late summer (March 2016), winter (July 2016), and early summer (December 2016) for study sites, going from upstream to downstream. The top right shows mean daily flow volumes and seston concentrations from October 2015 to September 2016

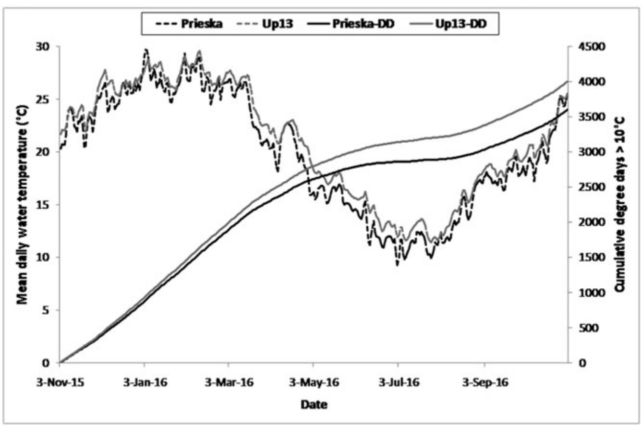

FIGURE 4 Mean daily water temperatures for the cooler, upstream site at Prieska versus the warmer downstream site at Upington (Up13). Cumulative degree days > 10°C are indicated on the second axis

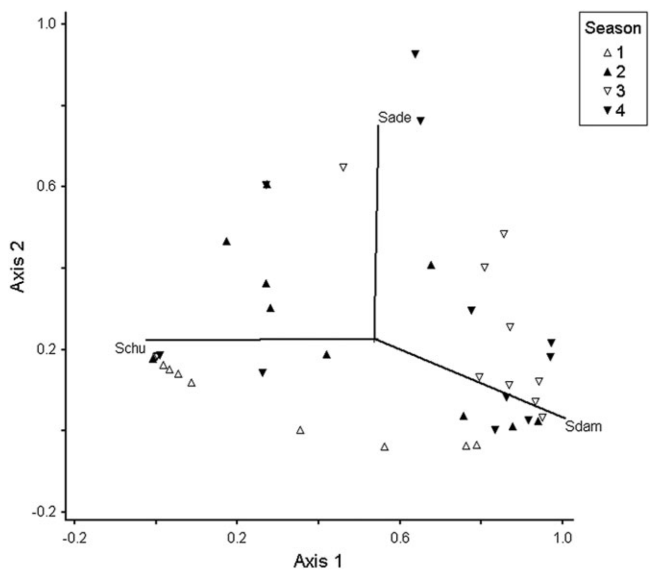

FIGURE 5 Bray–Curtis ordination of sites surveyed, based on simuliid species data (percent contribution of each species' relative abundance to total number per sample); season codes: 1 = late spring; 2 = late summer; 3 = winter; 4 = early summer (after Rivers-Moore & Palmer, 2018). Sade = *Simulium adersi*; Schu = *Simulium chutteri*; Sdam = *Simulium damnosum*

What was apparent, however, was a seasonal shift in water quality, with winter water quality driven by higher conductivity and clarity values (Figure 7; Table 3). Sites were generally clearer but with higher conductivity levels and higher alkalinity during winter and transitioning to more turbid, neutral pH conditions with greater dilution of salts in summer. Conversely, sites could be divided into three distinct thermal groups, based on their thermal metrics (Figure 8). Here, the Douglas and Prieska sites (Sites 1–2) clustered together as cooler sites than the remainder of the main channel Orange River sites, which showed greater thermal homogeneity.

Node probabilities were calculated for the cooler and warmer site groups based on 31 and 407 records of system states, respectively.

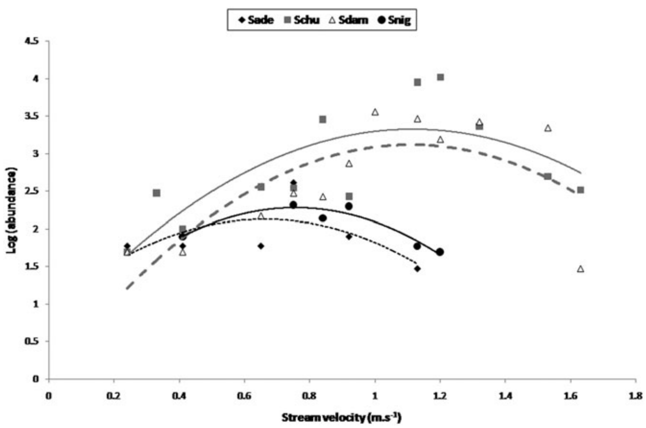

FIGURE 6 Log-transformed abundances of four species of blackfly based on combined sample data from November 2015 and March 2016. Velocity preference curves are based on maximum abundances for velocity values (see Table 3). Sade = *Simulium adersi*; Schu = *Simulium chutteri*; Sdam = *Simulium damnosum*; Snig = *Simulium nigritarse*

TABLE 2 Second-order polynomial equations describing relationship between stream velocity and relative abundance for four species of *Simulium* (see Figure 6)

Species	Equation	R^2
Simulium adersi	$y = -2.712x^2 + 3.589x + 0.942$	0.41
Simulium chutteri	$y = -2.192x^2 + 4.886x + 0.602$	0.59
Simulium damnosum	$y = -2.522x^2 + 5.600x + 0.008$	0.58
Simulium nigritarse	$y = -3.150x^2 + 4.766x + 0.479$	0.92

Outbreak probabilities were four times higher in the upper zone than in the lower zone (27.1% and 5.7%, respectively). Node sensitivity relative to the outbreak node for each model showed different combinations of variables having greater leverage for each BN (Table 4). Relative sensitivities illustrated that abiotic drivers were more than eight times more influential than biotic conditions in affecting outbreak probabilities, and with channel type and seston concentration being more important to model accuracy than water temperature and reed abundance. Model predictions for the outbreak probability node indicated <2% and 15% inaccuracy for the lower and upper zone models, respectively, than did the test case file data. Probabilities for the parent nodes compared favourably with the probabilities derived from return intervals on time series (Table 5), with the exception of seston concentrations. Here, probabilities based on the case files for turbid conditions were two thirds of what would be expected from the time series data. An example of the conditional probabilities for "low" versus "high" outbreak probability conditional upon three parent nodes is provided in Table 6.

TABLE 3 Eigenvalues for seasonal water quality variables associated with the 11 blackfly study sites on the middle and lower Orange River

	PC Axis 1	PC Axis 2
	Cumulative % variance	
	50.28	76.56
Variable	Eigenvalues	
pH	−0.365	−0.687
Turbidity	0.571	−0.373
Conductivity	0.574	−0.428
Season	0.460	0.453

Note. PC: principle component.

The final BN model represented relationships among all system variables, with probabilities of system states changing in accordance to the understood system behaviour and relationships (Figure 9). Eleven nodes were identified for the BN model, with each node having two to three states. Data indicated an increase in the probability of pest blackfly outbreaks between preimpoundment and postimpoundment flows. The highest probabilities of outbreaks, according to the model, were for February to April (Figure 10). Verification data, using maximum monthly values of the fly worry index, showed a lag of 1 to 2 months.

4 | DISCUSSION

Water quality conditions were relatively consistent across the 600-km study axis, with the exception of peripheral habitat that caters for different blackfly species with specific water quality preferences.

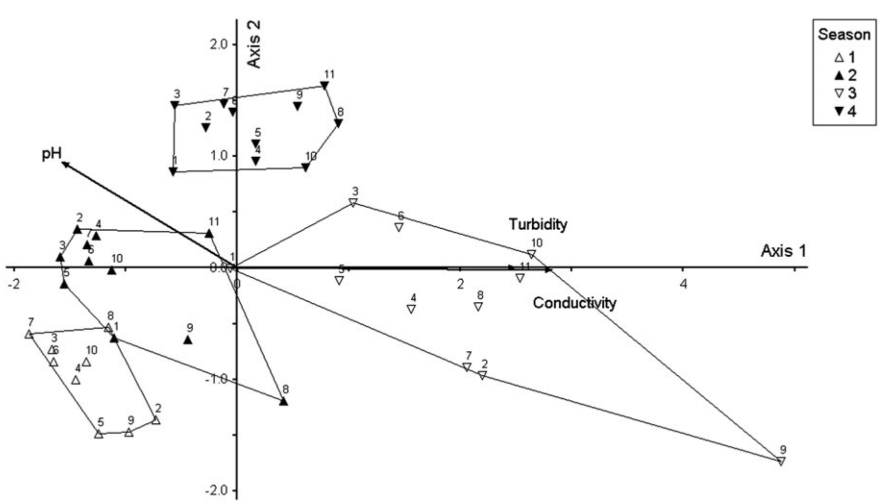

FIGURE 7 Principal component analysis of study sites based on water quality variables for November 2015 and March, July, and December 2016. Seasons 1–4 represent the data collected for the late spring (November 2015), late summer (March 2016), winter (July 2016), and early summer (December 2016) sampling periods, respectively

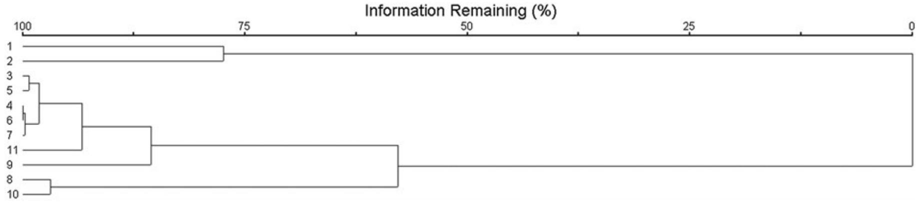

FIGURE 8 Cluster classification of study sites based on water temperature metrics. Sites were assigned to a cooler water temperature group (Sites 1–2) and a warmer water temperature group (Sites 3–10) according to the 75% information remaining threshold

TABLE 4 Node sensitivity relative to the "outbreak probability" node

Node	Lower	Upper
Outbreak probability	100.0	100.0
Simulium pest complex	19.5	17.4
Abiotic	14.0	14.3
Channel type	10.8	0.1
Seston concentration	10.0	14.5
Discharge	2.1	0.0
Biotic	1.6	0.3
Benthic algae	1.2	0.4
Reeds	0.2	0.0
Larvicidal efficacy	0.1	22.5
Water temperature	0.1	0.4

TABLE 5 Parent node state probabilities

Node	State probability	Case probability
Channel type	20% multiple/ 80% single	23.72% multiple/ 76.28% single
Discharge	37% < 100 $m^3 \cdot s^{-1}$/ 63% > 100 $m^3 \cdot s^{-1}$	36.19% < 100 $m^3 \cdot s^{-1}$/ 63.81% > 100 $m^3 \cdot s^{-1}$
Seston concentration	42% < 60 $mg \cdot L^{-1}$/ 58% > 60 $mg \cdot L^{-1}$	60.64% < 60 $mg \cdot L^{-1}$/ 39.36% > 60 $mg \cdot L^{-1}$
Water temperature	30% < 20 °C/70% > 20 °C	31.79% < 20 °C/ 68.21% > 20 °C

TABLE 6 Conditional probability values for the "outbreak probability" node, based on three input nodes

Simulium pest complex	Water temperature	Larvicidal efficacy	Outbreak probability = low	Outbreak probability = high
Minor	Cool	Optimal	50	50
Minor	Cool	Suboptimal	99.22	0.78
Minor	Warm	Optimal	98.15	1.85
Minor	Warm	Suboptimal	99.24	0.76
Major	Cool	Optimal	50	50
Major	Cool	Suboptimal	75	25
Major	Warm	Optimal	98.88	1.12
Major	Warm	Suboptimal	8.33	91.67

Larvae of *S. chutteri* exhibit a wide tolerance of water quality conditions (conductivities of 2–55 $mS \cdot m^{-1}$). Similarly, de Moor (1982) noted that fluctuations in pH in the Vaal River were minor (7.8–8.4) and not considered to account for larval size variation. Water temperatures are favourable throughout the year for blackfly life history development, although the marked cooling during autumn and winter is likely to lead to reduced numbers of generations over this period, and favouring larger larvae that develop into more fecund adults. This is particularly so for the Prieska and Douglas sites, which are slightly cooler than the other downstream sites. However, an important difference between the Douglas and Prieska sites was that the former site experiences significantly higher levels of subdaily flow variability than do sites further downstream, as a consequence of the water releases from Van Der Kloof Dam for hydro-electric power generation. Although these constant high and low flow pulses over a 24-hr time period are mitigated with downstream distance, the ecological consequence is that pest blackfly are more abundant at elevated but stable flows and less abundant at elevated by highly variable subdaily flows.

The BNs successfully incorporated the interactions of five environmental variables underpinning system switching between blackfly species complexes, together with the compounding effects of larvicidal efficacy on outbreak probability. Prediction accuracy when compared with test data was good, although the poorer performance for the upper zone illustrates the importance of understanding spatial modelling domains, and the need for two BNs in our system to cater for the upper, cooler sites and the warmer, lower sites. The model further demonstrated the considerable increase in outbreak probabilities postimpoundment and the most critical months where control actions need to be focussed. The probabilities derived for the parent nodes from the case files performed well against the probabilities derived from long-term time series, with the exception of seston concentration, which underrepresented ambient conditions. This has been shown to be a major driver underpinning system switches between different species complexes of blackfly (Rivers-Moore & Palmer, 2018) and would require long-term monitoring data to validate the actual probabilities. There was poor agreement between monthly outbreak probabilities and historical outbreak data. However, the time lag between the outbreak probabilities and the fly worry index scores is not surprising given that the BN model prediction probability of suitable habitat conditions for larvae is instantaneous, whereas the fly worry index reflects adult blackfly: Depending on water temperatures, the time required for life cycle completion is 12–24 days (de Moor, 1989).

Although impetus of the control programme may fade during periods when the problem has "gone away," this has not reduced the

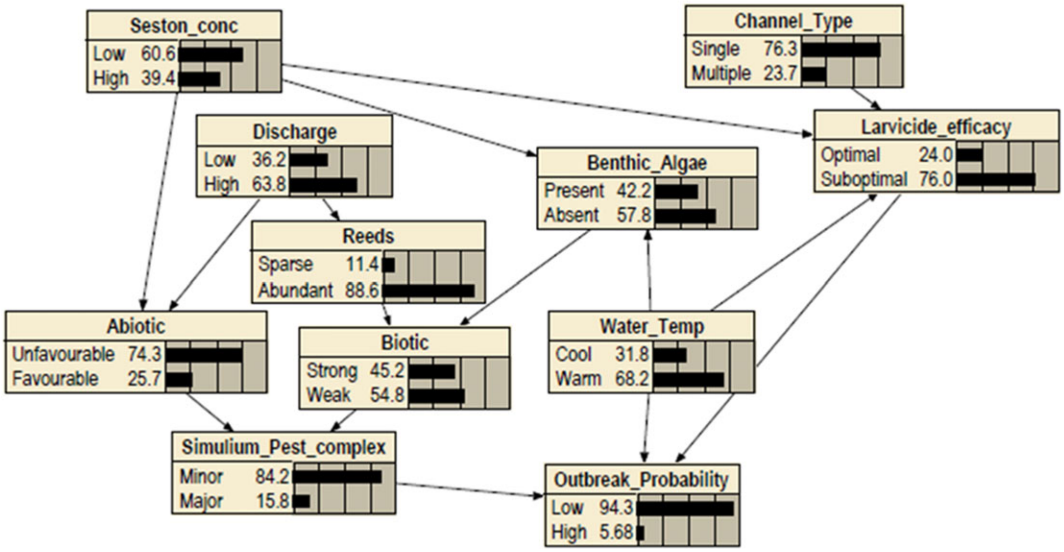

FIGURE 9 Bayesian network model showing parent and child nodes together with management and utility nodes [Colour figure can be viewed at wileyonlinelibrary.com]

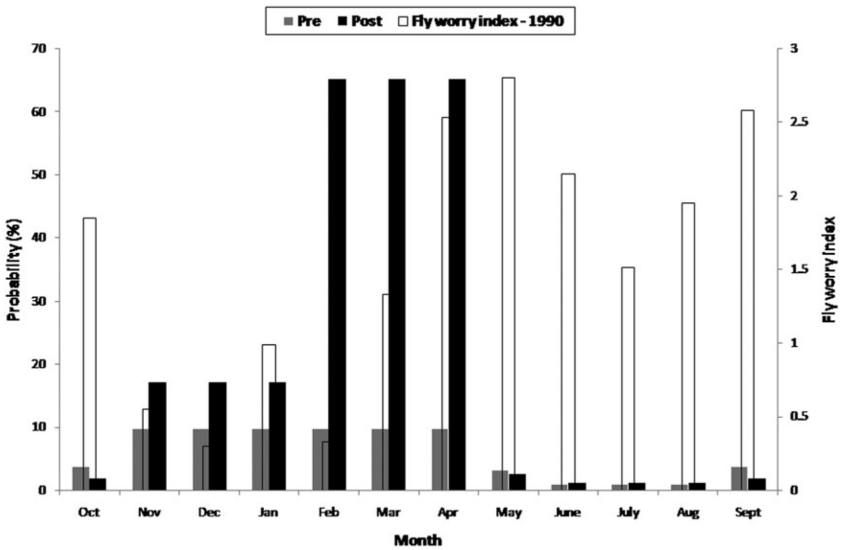

FIGURE 10 Seasonal variation in blackfly outbreak probabilities for preimpoundment and postimpoundment flow conditions

need for a stakeholder-driven, holistic, and proactive longer term solution to the problem. Ultimate solutions to this problem are, however, constrained by the conflicting resource needs of the stakeholder sectors along the middle and lower Orange River. Thus, flow manipulation may be feasible in theory, but its application is complicated by the income that would be lost through hydro-electric power generation in winter months, where power demand is highest. Agricultural activities are typically mixed, with the same land owners who suffer livestock losses also requiring irrigation water for vineyards. The lower and middle Orange River blackfly problem can truly be described as a "wicked problem," defined as difficult or impossible to solve because of contradictory user requirements (Rittel & Webber, 1973). In such situations, the challenge is to integrate the confounding effects of multiple variables into an integrated system understanding.

This model presents a framework for refining the current monitoring programme, where field visits that already score blackfly larval densities should be used as opportunities to record turbidity, presence/absence of benthic algae, blackfly species, and periodic downloading of water temperatures form data loggers. It is simple enough to be readily altered and adapted as needs dictate, while providing a simple template that can be applied across different thermal zones. Such model traits are regarded as desirable for uptake of BN models (Lynam, Drewry, Higham, & Mitchell, 2010). In terms of variables to be monitored, the node sensitivities in the predictive model highlight that *Simulium* pest species and larval densities plus water clarity are critical to be monitored, whereas water temperature and presence of benthic algae are desirable but less important to monitor. Even with limited data, BNs are regarded as being robust and able to show good prediction accuracy (Uusitalo, 2007). Ongoing monitoring data of a few easily measured variables, if incorporated as additional records into expanded case files, will serve to improve prediction accuracy. This could readily happen through the uploading of data onto an online platform such as a mobile phone app, with information periodically collated and analysed by a BN administrator. The inclusion of cost–benefit utility nodes together with a management node of alternative management interventions would extend the current BN into a decision network where alternative scenarios may be objectively evaluated.

Given that the environmental variables driving pest blackfly outbreaks in affected river systems globally are similar, this approach has potential to be applied elsewhere. With blackfly species being ubiquitous across global aquatic ecosystems, and conforming to the labral fan-type classification of Palmer and Craig (2000), the model could accommodate alternative blackfly species classified as “major” and “minor” pest species with relative ease. This would only require collection of data for the variables listed above, with relevant return interval probabilities calculated. We conclude that this study has potential for consolidating much of the previous research related to the Blackfly Control Programme into a useful predictive management framework. As the accuracy of predictions improves as the number of case files increases through the software learning algorithms, this model will become increasingly useful if updated. Together, the components of this framework are different from previous aspects of the Blackfly Control Programme, because they provide a structured means for auditing the successes and failures of the Blackfly Control Programme, and there is the basis for evaluating the most likely scenarios of future blackfly outbreaks in response to climate change-induced water temperature increases.

ACKNOWLEDGEMENTS

We thank the Water Research Commission (K5/2459) and Red Meat Research & Development South Africa for funding and supporting this research. Mnr. Hoffie Joubert from Agri-Noord Kaap and the staff from the Department of Agriculture, Forestry and Fisheries are thanked for their guidance on field sampling sites. Robert Palmer is thanked for comments on the Bayesian network; Sashin Naidoo, Esther Ndou and Jarryd Gilham are thanked for their assistance on the project, specifically for field work and processing of samples. The two reviewers are sincerely thanked for their inputs and comments.

ORCID

Nicholas A. Rivers-Moore http://orcid.org/0000-0002-6546-4215
Trevor R. Hill http://orcid.org/0000-0001-7015-6906

REFERENCES

Batchelor, C., & Cain, J. (1999). Application of belief networks to water management studies. *Agricultural Water Management*, *40*, 51–57.

Bonkewitzz, A. N., & Palmer, R. W. (1997). BFControl© version 1.0. In: Palmer, R.W. *Principles of integrated control of blackflies (Diptera: Simuliidae) in South Africa*. WRC Report No 650/1/97, Water Research Commission, Pretoria, South Africa.

Cain, J. (2001). Planning improvements in natural resources management: Guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond. Centre for Ecology and Hydrology, Wallingford.

Chalifour, A., Boisvert, J., & Back, C. (1990). Optimization of insecticide treatments in rivers: Application of graph theory for planning a black fly control program. *Canadian Journal of Fisheries and Aquatic Sciences*, *47*, 2049–2056.

Crosskey, R. W. (1973). Chapter 3b. Simuliidae (Black-flies, German: Kriebelmucken). In K. G. V. Smith (Ed.), *Insects and other arthropods of medical importance*. London: British Museum (Natural History).

de Moor, F. C. (1982). Determination of the number of instars and size variation in the larvae and pupae of *Simulium chutteri* Lewis 1965 (Diptera: Simuliidae) and some possible bionomical implications. *Canadian Journal of Zoology*, *60*, 1374–1382.

de Moor, F. C. (1989). Alternative life-history styles in Simuliidae (Insecta, Diptera). In *Alternative life-history styles of animals* (pp. 293–316). Netherlands: Springer.

de Moor, F. C. (1994). Aspects of the life history of *Simulium chutteri* and *S. bovis* (Diptera: Simuliidae) in relation to changing environmental conditions in South African rivers. *Verhandlungen - Internationale Vereinigung fur Theoretische und Angewandte Limnologie*, *25*, 1817–1821.

de Moor, F. C. (2003). Chapter 5: Simuliidae. In: JA Day, AD Harrison, IJ de Moor. 2003. *Guides to the freshwater invertebrates of Southern Africa, vol. 9: Diptera*. WRC Report No. TT201/02/Water Research Commission, Pretoria, South Africa.

Fortin, M.-J., & Dale, M. (2005). *Spatial analysis: A guide for ecologists*. Cambridge: Cambridge University Press.

Fredeen, F. J. H. (1977). Some recent changes in black fly populations in the Saskatchewan River system in western Canada coinciding with the development of reservoirs. *Canadian Water Resources Journal*, *2*, 90–101.

Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs* (2nd ed.). New York: Springer.

Kjaerulff, U. B., & Madsen, A. L. (2008). *Bayesian networks and influence diagrams: A guide to construction and analysis*. New York: Springer.

Lynam, T., Drewry, J., Higham, W., & Mitchell, C. (2010). Adaptive modelling for adaptive water quality management in the Great Barrier Reef region, Australia. *Environmental Modelling and Software*, *25*, 1291–1301.

McCune, B., & Mefford, M. J. (2011). *PC-ORD. Multivariate analysis of ecological data. Version 6.16*. Gleneden Beach, OR: MjM Software.

Mullins, W. (2007). Chapter 7: Cost–benefit analysis. In: Palmer RW, Rivers-Moore NA, Mullins W, McPherson V, Hattingh L. 2007. *Guidelines for integrated control of pest blackflies along the Orange River*. WRC Report No. 1558/1/07, Water Research Commission, Pretoria, South Africa.

Norsys Software Corporation (2010). Netica v. 4.16. URL: http://norsys.com

O'Keeffe, J. H., & de Moor, F. C. (1988). Changes in the physico-chemistry and benthic invertebrates of the Great Fish River, South Africa, following an interbasin transfer of water. *Regulated Rivers: Research and Management*, *2*, 39–55.

Palmer, R. W. (1994). A rapid method of estimating the abundance of immature blackflies (Diptera: Simuliidae). *Onderstepoort Journal of Veterinary Research*, *61*, 117–126.

Palmer, R. W. (1997). Principles of integrated control of blackflies (Diptera: Simuliidae) in South Africa. WRC Report No. 650/1/97, Water Research Commission, Pretoria, South Africa.

Palmer, R. W., & Craig, D. A. (2000). An ecological classification of primary labral fans of filter-feeding black fly (Diptera: Simuliidae) larvae. *Canadian Journal of Zoology*, *78*, 199–218.

Palmer, R. W., & O'Keeffe, J. H. (1995). Distribution and abundance of blackflies (Diptera: Simuliidae) in relation to impoundments in the Buffalo River, Eastern Cape, South Africa. *Freshwater Biology*, *33*, 109–118.

Palmer, R. W., & Palmer, A. R. (1995). Impacts of repeated applications of *Bacillus thuringiensis* var. *israelensis* de Barjac and temephos, used in blackfly (Diptera: Simuliidae) control, on macroinvertebrates in the middle Orange River, South Africa. *Southern African Journal of Aquatic Sciences*, *21*, 35–55.

Palmer, R. W., & Rivers-Moore, N. A. (2008). Evaluation of larvicides in developing management guidelines for long-term control of pest blackflies along the Orange River. *Onderstepoort Journal of Veterinary Research*, *75*, 299–314.

Palmer, R. W., Rivers-Moore, N. A., Mullins, W., McPherson, V., & Hattingh, L. (2007). Guidelines for integrated control of pest blackflies along the Orange River.WRC Report No. 1558/1/07, Water Research Commission, Pretoria, South Africa.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*, 155–169.

Rivers-Moore, N. A., Bangay, S., & Palmer, R. W. (2008). Optimization of *Bacillus thuringiensis* var. *israelensis* (VectoBac®) applications for the Blackfly Control Programme on the Orange River, South Africa. *Water SA*, *34*, 193–198.

Rivers-Moore, N. A., Dallas, H. F., & Morris, C. (2013). Towards setting environmental water temperature guidelines: A South African example. *Journal of Environmental Management*, *82*, 39–50.

Rivers-Moore, N. A., Dallas, H. F., & Ross-Gillespie, V. (2013). Life history does matter in assessing potential impacts of thermal changes on aquatic macroinvertebrates. *River Research and Application*, *29*, 1100–1109.

Rivers-Moore, N. A., de Moor, F. C., Birkholz, S. A., & Palmer, R. W. (2006). Estimation of preferred water flow parameters for four species of *Simulium* (Diptera: Simuliidae) in small clear streams in South Africa. *African Journal of Aquatic Science*, *31*, 261–269.

Rivers-Moore, N. A., de Moor, F. C., Morris, C., & O'Keeffe, J. (2007). Effect of flow variability modification and hydraulics on invertebrate communities in the Great Fish River (Eastern Cape Province, South Africa), with particular reference to critical hydraulic thresholds limiting larval densities of *Simulium chutteri* Lewis (Diptera, Simuliidae). *River Research and Applications*, *23*, 201–222.

Rivers-Moore, N. A., Hughes, D. A., & de Moor, F. C. (2008). A model to predict outbreak periods of the pest blackfly *Simulium chutteri* Lewis (Simuliidae, Diptera) in the Great Fish River, Eastern Cape Province, South Africa. *River Research and Applications*, *24*, 132–147.

Rivers-Moore, N. A., & Palmer, R. W. (2018). The influence of turbidity and water temperature on black fly species in the middle and lower Orange River, South Africa. *Canadian Journal of Zoology*, *96*, 614–621.

Rivers-Moore, N. A., Palmer, R. W., & Dallas, H. F. (2014). Assessing the relative culpability of *Simulium* (Diptera: Simuliidae) species in recent blackfly outbreaks along the middle Orange River, South Africa. *Canadian Journal of Zoology*, *92*, 505–513.

Rodionov, S. (2006). *Sequential regime shift detection v3.2*. Seattle: Pacific Marine Environmental Laboratory, NOAA/PMEL.

Rosenberg, M. S., & Anderson, C. D. (2011). PASSaGE: Pattern analysis, spatial statistics and geographic exegesis. Version 2. *Methods in Ecology and Evolution*, *2*, 229–232.

Sheldon, A. L., & Oswood, M. W. (1977). Blackfly (Diptera: Simuliidae) abundance in a lake outlet: Test of a predictive model. *Hydrobiologia*, *56*, 113–120.

Snaddon, C. D., & Davies, B. R. (1998). A preliminary assessment of the effects of a small South African inter-basin water transfer on discharge and invertebrate community structure. *Regulated Rivers: Research and Management*, *14*, 421–441.

Stewart-Koster, B., Bunn, S. E., Mackay, S. J., Poff, N. L., Naiman, R. J., & Lake, P. S. (2010). The use of Bayesian Networks to guide investments in flow and catchment restoration for impaired river ecosystems. *Freshwater Biology*, *55*, 243–260.

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, *203*, 312–318.

How to cite this article: Rivers-Moore NA, Hill TR. A predictive management tool for blackfly outbreaks on the Orange River, South Africa. *River Res Applic*. 2018;34:1197–1207. https://doi.org/10.1002/rra.3357