Modelling a pike (*Esox lucius*) population in a lowland river using a cellular automaton

I.S. Pauwels [a](#_bookmark0), [b](#_bookmark1),[⁎](#_bookmark3), A.M. Mouton [a](#_bookmark0), [b](#_bookmark1), J.M. Baetens [c](#_bookmark2), S. Van Nieuland [a](#_bookmark0), B. De Baets [c](#_bookmark2), P.L.M. Goethals [a](#_bookmark0)

a *Ghent University, Laboratory of Environmental Toxicology and Aquatic Ecology, J. Plateaustraat 22, B-9000 Ghent, Belgium*

b *Research Institute for Nature and Forest (INBO), Kliniekstraat 25, B-1070 Brussels, Belgium*

c *KERMIT: Research Unit Knowledge-based Systems, Ghent University, Coupure links 653, B-9000 Ghent, Belgium*

*Keywords:*

Cellular automaton Fish

Pike

Habitat suitability Ecological knowledge Spatio-temporal model

a b s t r a c t

Cellular automata (CAs) allow for transparent modelling of complex systems based on simple transition rules and are ﬂexible in incorporating individual differences and local interactions. They may therefore be partic- ularly suited to answer river management questions that could not be addressed by existing habitat suitabil- ity models, such as the optimal distance between spawning grounds. This study explores the usability of CAs for spatio-temporal modelling of a pike population to support river management. Speciﬁcally, we evaluated the usability of the CA model by analyzing its sensitivity to three model parameters: the number of pike in the grid, the initial pike distribution and the grid resolution. The model includes habitat characteristics and basic expert knowledge on the ecology of pike and was tested on a 10 km stretch of the river Yser in Flanders (Belgium). Simulation results showed that the model converged to a realistic pike distribution over the study area only at high pike density and low grid resolution, irrespective of the initial pike distribution. Pike density and grid resolution affected the sensitivity to the initial pike distribution in the grid. Speciﬁcally, the sensitivity was high at low pike density and high grid resolution, and absent when pike density was high. This analysis indicated that initial conditions and cell size may have a severe impact on the model output, illustrating the importance of ﬁrstly analyzing this impact before conducting further analyses. Depending on the outcome of such analyses, CAs can be a promising modelling technique to evaluate and predict the effect of river resto- ration on pike populations.

1. Introduction

Sampling campaigns in Flanders revealed a poor condition of pike populations in rivers since 1950 ([www.vis.milieuinfo.be](#_bookmark33)). In 1949, pike was widespread, but population viability decreased in all rivers and by the mid 1970s pike was no longer observed. Due to the improvement of the water quality, pike was observed again since 1990, although signiﬁcantly less abundant than before ([De Nayer and](#_bookmark18) [Belpaire, 1997; Goethals et al., 2006; Vandenabeele et al., 1998](#_bookmark18)). In an attempt to rehabilitate pike populations, reintroduction programmes were started. These programmes were only moderately successful due to the poor water quality, the loss of suitable habitat ([Maeckelberghe,](#_bookmark18) [2002; Vandenabeele et al., 1998](#_bookmark18)), and the obstruction of ﬁsh migration ([Coeck, 2002; Knaepkens et al., 2004](#_bookmark18)). Insights into the spatio-temporal dynamics of pike populations are thus essential to better predict the impact of pike introduction and conservation actions. Speciﬁcally, a successful, integrated management not only requires information on

\* Corresponding author at: Ghent University, Laboratory of Environmental Toxicolo- gy and Aquatic Ecology, J. Plateaustraat 22, B-9000 Ghent, Belgium. Tel : + 32 9 264 37 76; fax: + 32 9 264 41 99.

*E-mail address:* [Ine.Pauwels@ugent.be](mailto:Ine.Pauwels@ugent.be) (IS. Pauwels).

the speciﬁc habitat requirements, but also on the spatial distribution of the suitable habitats and on the migration of pike between these habitats during its life cycle.

Although species distribution models may reveal new insights into the ecology of pike and their functioning within ecosystems ([Glasbergen, 2001; Guisan and Zimmermann, 2000; Inskip, 1982; Kerle](#_bookmark24) [et al., 2001; Mouton, 2008; Zarkami, 2008](#_bookmark24)), they are often limited to either temporal or spatial ﬂuctuations in pike distribution. Consequent- ly they do not provide information about the spatial dynamics of pike, although this species typically needs different, spatially separated habitats to successfully complete its life cycle. Geographical Informa- tion Systems (GIS) and mathematical paradigms for spatially explicit and dynamic modelling such as partial differential equations (PDEs), individual-based models (IBMs) and cellular automata (CAs) were introduced in ecological modelling to overcome this problem. Recently, [Chen et al. (2011)](#_bookmark18) reviewed these modelling paradigms, their usability and shortcomings. They point out that, although GIS and remote sens- ing are powerful tools for spatial analysis, these give an inherently static view of the world and are unable to capture and model dynamic pro- cesses. PDEs can describe the abundance of a species both spatially and temporally, but neglect spatial heterogeneity, local interactions and individual differences, since they rely on mean-ﬁeld assumptions

([Chen and Mynett, 2006](#_bookmark18)). In contrast, IBMs allow to model the dynamics of individuals in an ecosystem that shows non-uniform variation of environmental variables ([Perry and Enright, 2007](#_bookmark25)), and therefore have been applied frequently in this context ([Le Fur and Simon, 2009;](#_bookmark18) [Li et al., 2010; Uchmanski et al., 1999; Yemane et al., 2009](#_bookmark18)). An IBM con- siders each ﬁsh as an individual of which the condition and position are tracked in time based on characteristics such as gender, age and a limit- ed set of environmental variables. CA models are similar to IBMs but focus on the effect of spatially and temporally heterogeneous environ- mental conditions on the population, rather than on the individual characteristics and behaviour. Since river managers are more interested in the impact of environmental changes on entire populations than on individual ﬁshes, CA models may be more appropriate than IBMs to support river management and to reveal effective restoration measures. Therefore, this paper aims to evaluate the usability of CAs to describe the spatial dynamics of a pike population in a river. We simulated pike migration in relation to vegetation and bank structure in a 10 km stretch of the river Yser. A sensitivity analysis revealed the impact of three model parameters: the number of pike in the grid, the initial pike distribution and the grid resolution. We discussed the ability of CA to evaluate and predict the effect of river restoration on pike populations and identiﬁed opportunities for future research.

1. Material and methods
   1. *Study area and data collection*

The study area included a 10 km stretch of the 44 km Belgian part of the 76 km river Yser ([Fig. 1](#_bookmark4)). The river Yser drains an area of 1101 km² ([Heylen, 1997](#_bookmark18)) and is a typical precipitation river, with an average annu- al discharge of 1.44 m³s− 1 ([Provoost, 1997](#_bookmark26)), a peak ﬂow of 5.7 m³s− 1, and a base ﬂow of 0.8 m³s− 1 ([Heylen, 1997](#_bookmark18)). The drainage is inﬂuenced by a tidal sluice at the estuary that prevents tidal water level ﬂuctuation. Upstream of this sluice there are no migration barriers in the main river and pike can migrate freely between the river Yser and most of its

tributaries. Water depth is 2.5 m on average, although some speciﬁc habitats, such as foreshores and artiﬁcial spawning grounds are only

0.7 and 0.4 m deep on average ([Mouton et al., 2011](#_bookmark22)). Near the French border, the river is 8 to 10 m wide, whereas at the river mouth the width is already 25 m ([Mouton et al., 2009c](#_bookmark21)). The water quality of the river Yser and its tributaries is suitable for ﬁsh and is relatively constant since there are only a limited number of point pollution sources in the study area ([www.vmm.be](http://www.vmm.be/)).

The 10 km study area was divided in cells, each representing 500 and 10 m of the river stretch. At both banks of each cell, riparian veg- etation and bank structure were assessed visually. Vegetation was classiﬁed into four classes: reed (Phragmites australis Cav.), shrubs, scarce vegetation and no vegetation (Research Institute of Nature and Forest, unpublished data). Bank structure was classiﬁed into four classes: vertically reinforced banks, tiles, foreshores and semi- natural banks (Research Institute of Nature and Forest, unpublished data). Spawning grounds were considered a special bank type and assessed differently. They are stagnant but connected to the river channel. Depth is variable, while bank vegetation is dominated by tall herbs and reed. Artiﬁcial foreshores consist of a row of 0.2 m diame- ter wooden posts that are positioned in the river channel between 0.7 m and 2.5 m from the river bank. The distance between two consecutive posts varies between 0.05 m and 0.2 m, such that these posts protect the riparian habitats from shipping wave action. Vegetation in the zone between the posts and the river bank consists mainly of reed and vege- tation cover ranges from 0% to 100%. Semi-natural habitats are highly heterogeneous and characterised by varied bank vegetation such as trees, herbs and submersed vegetation ([Mouton et al., 2011](#_bookmark22)). Along the Belgian reach of the river Yser, 43% of the banks are artiﬁcially embanked and straightened with concrete or submersed bricks (INBO, unpublished data). The foreshores comprise 12% of the total length, the spawning grounds 3% and the remaining banks are deﬁned as being semi-natural ([Mouton et al., 2009c, 2011](#_bookmark21)). Pike habitat suitability was quantiﬁed based on vegetation and bank structure data by assigning a habitat suitability index (HSI; [U.S. Fish and Wildlife Service, 1981](#_bookmark31)) between

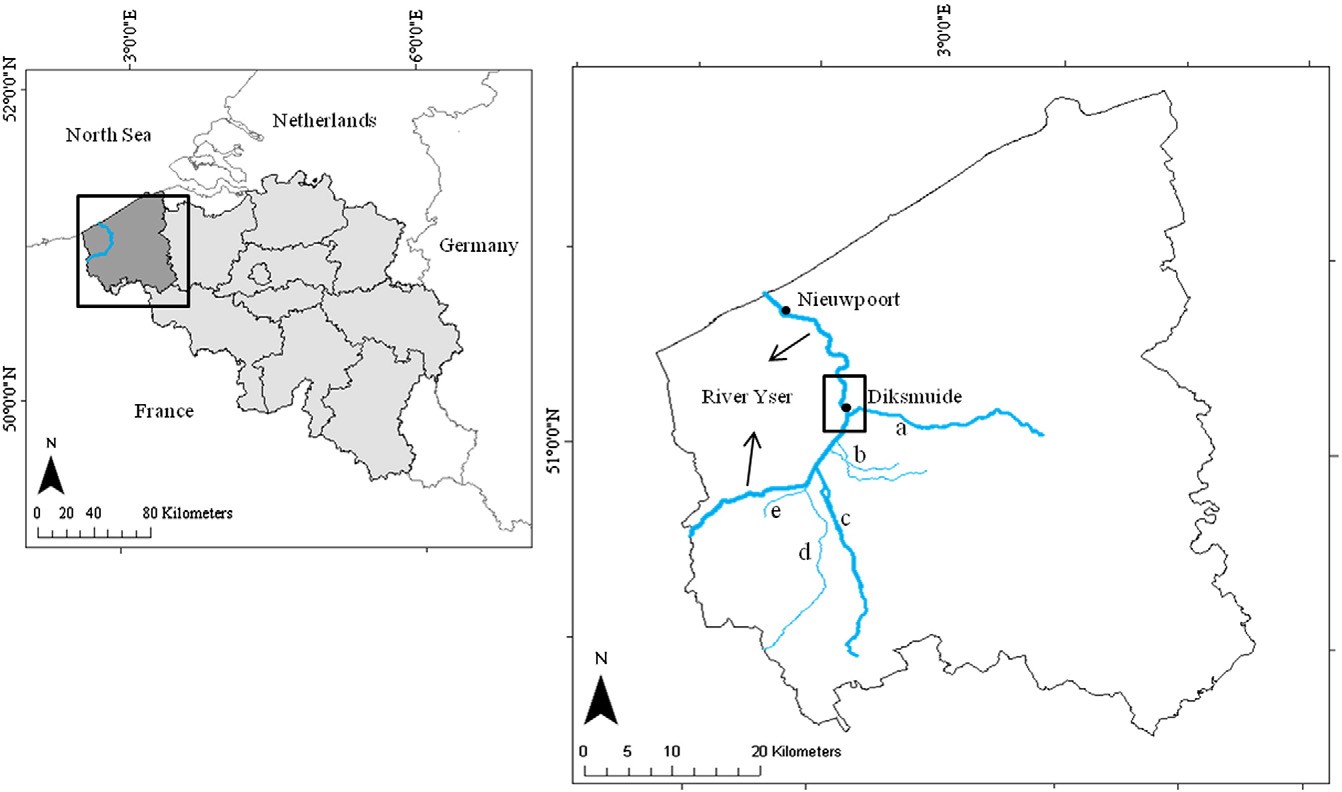


Fig. 1. Location of the 10 km study area (box) in the Belgian part of the river Yser, located in Flanders, Belgium (a = Handzamevaart, b = Houtensluisvaart and Stenensluisvaart, c = Canal from the city of Ieper to the river Yser, d = Kemmelbeek, e = Boezingegracht).

Table 1

Deﬁnition of the Habitat Suitability Index (HSI) of each cell bank in the study area based on the HSIs assigned to the different riparian vegetation and bank structure types. Only combinations of vegetation and bank structure types occurring in the study area were considered. Artiﬁcial spawning grounds were considered a particu- larly suitable habitat for pike and therefore a cell was assigned a HSI of 1 if it con- tained an artiﬁcial spawning ground.

|  |  |  |
| --- | --- | --- |
| Vegetation type (HSI) | Bank structure (HSI) | Cell bank HSI |
| Reed (0.8) | Semi-natural bank (0.8) | 0.8 |
| Shrub (0.6) | Semi-natural bank (0.8) | 0.7 |
| Reed (0.8) | Foreshores (0.6) | 0.7 |
| Shrub (0.6) | Tiles (0.6) | 0.6 |
| Scarce vegetation (0.4) | Semi-natural bank (0.8) | 0.6 |
| Scarce vegetation (0.4) | Tiles (0.6) | 0.5 |
| No vegetation (0.3) | Vertically reinforced bank (0.3) | 0.3 |

0.3 and 1 to each vegetation and bank structure class ([Table 1](#_bookmark5)). The HSI of a cell was then calculated as the average of the HSIs along the left and right bank, which were on their turn determined as the average of the HSIs of the vegetation and bank structure. Artiﬁcial spawning grounds were considered a particularly suitable habitat for pike and therefore a cell was assigned a HSI of 1 if it contains an artiﬁcial spawning ground.

This approach assumes that pike need shallow places to spawn and vegetation or other plant parts to hide during foraging ([Casselman and](#_bookmark18) [Lewis, 1996; Craig, 1996, 2008; Eklöv, 1997; Inskip, 1982](#_bookmark18)). Although other variables like temperature and oxygen concentration may also im- pact pike distribution ([Cooper et al., 2008; De Laak and van Emmerik,](#_bookmark18) [2006; Farrell et al., 2006; Jacobsen et al., 2007; Raat, 1988; Smith et al.,](#_bookmark18) [2007](#_bookmark18)), these variables were not taken into account because their spatial variation in the study area was too limited.

* 1. *Simulating pike migration with a CA*
     1. *Model structure and assumptions*

Characteristic of CAs is the discrete nature of the time, space and

i.e. *S*(*ci*, *t*). The CA grid in this study is 1-dimensional and consists of regular cells that have a state *S*(*ci*, *t*) which represents the pike abun- dance in a cell. At each discrete time step this cell state is computed based on speciﬁc transition rules, each of them depending on the state of the cell itself and the states of its neighbouring cells *cj* ([Chen et](#_bookmark18) [al., 2011; Wolfram, 1983](#_bookmark18)). The neighbouring cells are deﬁned by the neighbourhood, which was conﬁned to the nearest neighbours in this study.

The presented model is a stochastic CA in which all cells are updated simultaneously by the same transition rules. The impact of the HSI on pike migration is deﬁned by these transition rules ([Fig. 2](#_bookmark6)). The modelled population was assumed to consist only of adults and birth and death processes were not taken into account in the present study.

* + 1. *Transition rules*

The stochastic transition rules ([Minss et al., 2000](#_bookmark18)) quantify the number of pike that leave a given cell at each iteration step based on its HSI. The HSIs of the neighbouring cells determine the number of migrating pike that move from *ci* to *cj*, such that more pike will leave cell *ci* if the habitat of *ci* is less suitable. Previous research has shown that pike has speciﬁc habitat demands ([Casselman and Lewis,](#_bookmark18) [1996; Craig, 1996, 2008; Eklöv, 1997; Inskip, 1982](#_bookmark18)) and that they move to their preferred habitat if possible ([Haugen et al., 2006](#_bookmark32)). To in- tegrate random migration and the effect of other variables than habitat suitability, a stochastic component is added to the formula. Speciﬁcally, the number of pike in each cell, at a speciﬁc time step, is calculated based on its state at the previous time step *t* − 1, *S*(*ci*, *t* − 1) and on

the number of pike that emigrate from or immigrate to that cell

*E*(*ci*, *t* − 1) and *I*(*ci*, *t* − 1), respectively:

*S*ð*ci* ; *t*Þ ¼ *S*ð*ci*; *t*−1Þ−*E*ð*ci*; *t*−1Þ þ *I*ð*ci*; *t*−1Þ ð2:1Þ

The number of pike *E*(*ci*, *t* − 1) that emigrate from a cell *ci* at time step *t* − 1, is calculated as follows:

state domains. Speciﬁcally, a CA consists of an n-dimensional grid of (

( ( \ \\

regular or irregular cells *ci* that take a certain state s at time step t,

*E*ð*ci*; *t*−1Þ ¼ *R*ð*ci*; *t*−1Þ*S*ð*ci* ; *t*−1Þ ð1−*HSI*ð*ci* ÞÞ 1−*E*~

þ *E*~

ð2:2Þ

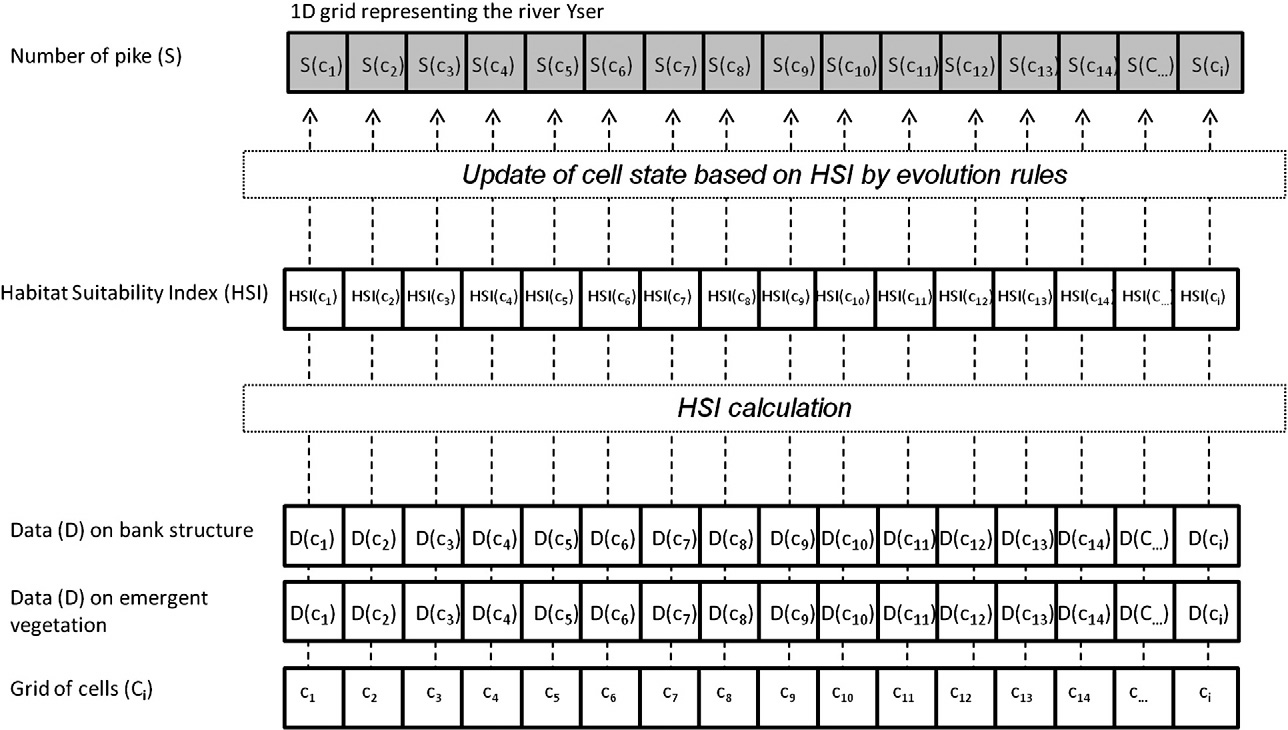


Fig. 2. Schematic overview of the model structure. Habitat Suitability Index (HSI) calculation was based on vegetation and bank structure data (D(c\_i)) of each grid cell c\_i. The cell state (S(c\_i)), the number of pike in each cell, was calculated at each time step by transition rules based on the habitat suitability in each cell (HSI(c\_i)).

Table 2

Overview of the 12 simulations evaluating the sensitivity of the model output to the total number of pike in the grid, initial pike distribution over the study area and grid resolution. The results of simulations FL and FM were not presented, because they were similar to simulations CL and CM.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cell length (representing | Number of pike in the grid | Initial pike distribution | Code | Result |
| length (m)) | (# pike·grid− 1) | (initial *H*) |  |  |
| 500 | 20 | Evenly distributed | CLE | [Fig. 5](#_bookmark10)A–D |
| 500 | 20 | All initially in one cell | CLU | [Fig. 6](#_bookmark11)A–D |
| 500 | 200 | Evenly distributed | CME | [Fig. 7](#_bookmark12)A–D |
| 500 | 200 | All initially in one cell | CMU | [Fig. 8](#_bookmark13)A–D |
| 500 | 1200 | Evenly distributed | CHE | [Fig. 9](#_bookmark14)A–D |
| 500 | 1200 | All initially in one cell | CHU | [Fig. 10](#_bookmark15)A–D |
| 10 | 1000 | All initially in one cell | FLE | Not shown |
| 10 | 1000 | Evenly distributed | FLU | Not shown |
| 10 | 10,000 | All initially in one cell | FME | Not shown |
| 10 | 10,000 | Evenly distributed | FMU | Not shown |
| 10 | 60,000 | Evenly distributed | FHE | [Fig. 11](#_bookmark16)A–D |
| 10 | 60,000 | All initially in one cell | FHU | [Fig. 12](#_bookmark17)A–D |

with *HSI*(*ci*) the HSI of cell *ci*, *E*~ a constant and *R*(*ci*, *t* − 1) the stochastic component of cell *ci* at iteration step *t* − 1. The latter terms ensure that there is always a random fraction of *S*(*ci*, *t* − 1) that emigrates,

irrespective of *HSI*(*ci*). This fraction is quantiﬁed by the constant *E*~

(which was set to 0.1 in all simulations), and by the stochastic com- ponent, *R*(*ci*, *t* − 1) ~ *U*(0.25, 0.5) (where U represents a uniform dis- tribution between 25 and 50% of the total number of pike).

Further, the number of pike that immigrated *I*(*ci*, *t* − 1) to *c* \_ *i* in time step *t* − 1, was calculated as follows:

with *E*(*cj*, *t* − 1) the number of pike leaving the neighbouring cell *cj* at time step *t* − 1, and |*N*(*ci*)| and |*N*(*ci*)| being the number of neighbours of respectively *ci* and *cj*, which was never higher than two, since the CA grid in this study was one dimensional.

* + 1. *Model output and sensitivity analysis*

By means of a sensitivity analysis the effect of three model charac- teristics was analyzed: the number of pike in the grid, the initial pike distribution and the grid resolution ([Table 2](#_bookmark7)). Six simulations were performed in a grid with 20 cells of 500 m long (a coarse grid; C), and six in a grid having 1000 cells of 10 m in length (a ﬁne grid; F). Each of the simulations differed in pike grid density and in the initial distribution, which was even (E) or uneven (U). At an even distribu- tion every cell of the grid contains an equal number of pike, whereas at an uneven distribution only the ﬁrst cell contains all the pike. Three different grid densities of pike were tested: 20 (L), 200 (M) and 1200

(H) pike·grid cell− 1 ([Table 2](#_bookmark7)). The tested densities were based on three

different ﬁeld observations of carrying capacity: 2.3–8.3 adults·ha− 1 by [Craig (1996)](#_bookmark18), 40–50 adults·ha− 1 and 1.1–1.6 adults·ha− 1 by [Margenau](#_bookmark18) [et al. (2008)](#_bookmark18). The simulations are identiﬁed with a three-letter code, in which the ﬁrst letter indicates the grid type (C or F), the second indicates the pike density (L, M or H) and the last indicates the initial distribution (E or U) ([Table 2](#_bookmark7)). Furthermore, in every simulation the spawning grounds were empty at the ﬁrst iteration step.

The different simulations were compared based on the number of iterations before model convergence and on the ﬁnal pike distribution obtained after convergence. This was assessed by measuring the net migration, which was deﬁned as the total change in number of pike of

j*N*ð*ci* Þj

0 1

*HSI* ð*ci* Þ ( \

all grid cells between two consecutive iteration steps. At model conver-

*I*ð*ci*; *t*−1Þ ¼ ∑*j*¼1 @

j*N*ð*cj* Þj

*E cj* ; *t*−1 A ð2:3Þ

gence the net migration ﬂuctuates around a steady-state value and no

∑*k*¼1 *HSI*ð*ck* Þ

longer affects the pike distribution across the study area.

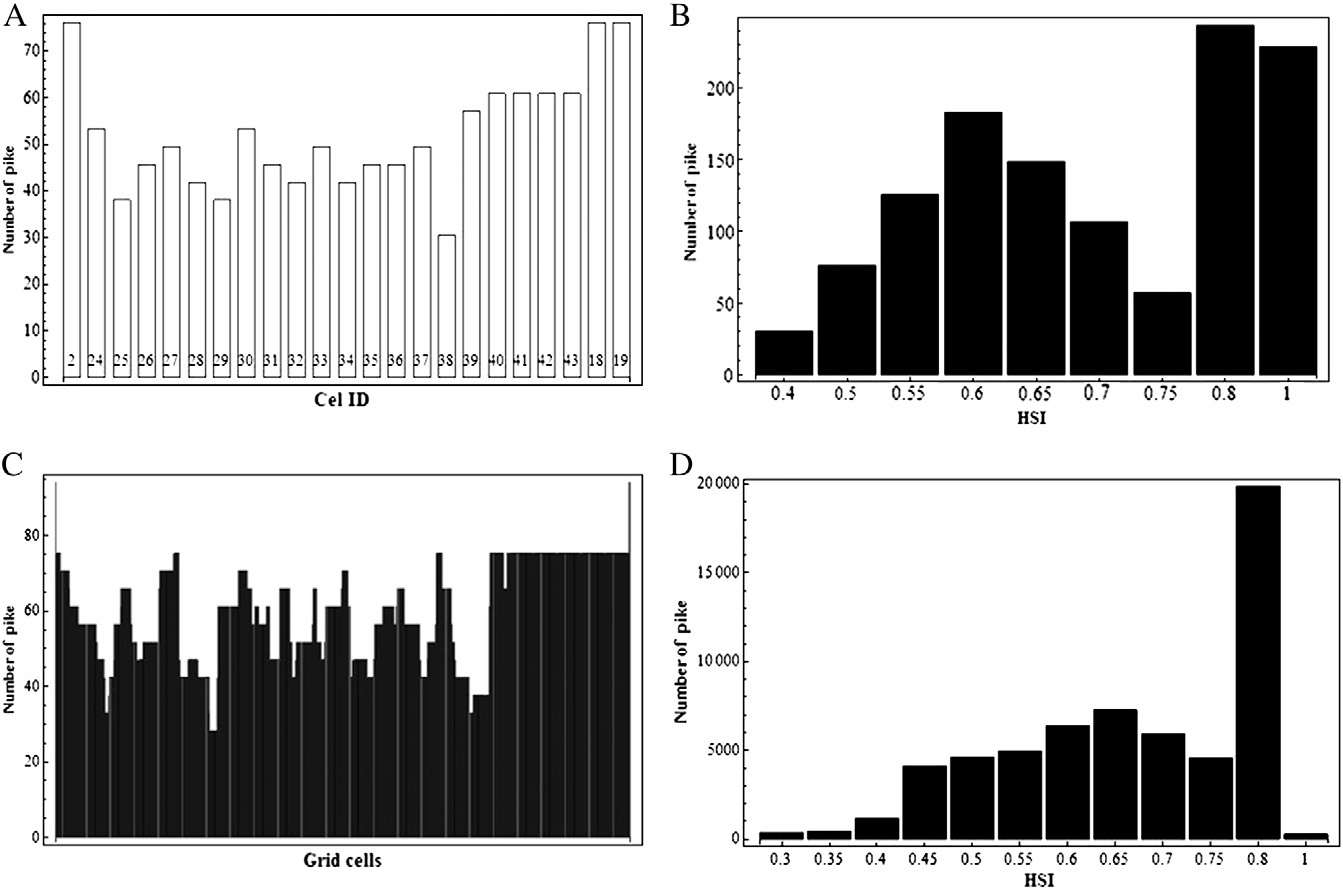


Fig. 3. Expected distributions of pike over the whole grid (A) and over the HSI classes (B) in the grid with cell lengths representing 500 m, and over the whole grid (C) and over the HSIs (D) in the grid with cell lengths representing 10 m.

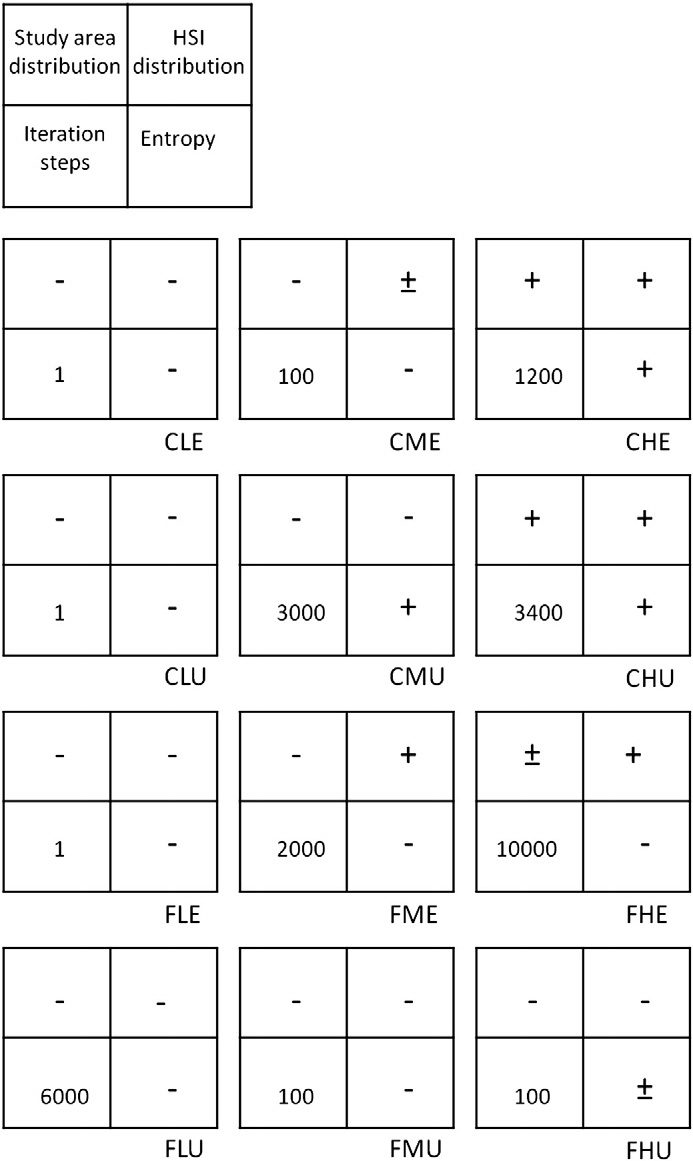


Fig. 4. Schematic overview of the simulation performances by good (+), moderate (±) or poor (-) ability to approximate the expected distribution over the study area (study area distribution, top left), over the HSIs (HSI distribution, top right) and en- tropy (bottom right), and the number of iteration steps needed before convergence (bottom left). The codes identify the simulation parameters: C: coarse grid, F: ﬁne grid, L: low pike density (1 pike.cell-1), M: medium pike density (10 pike.cell-1 ), H: high pike density (60 pike.cell-1), E: pike initially evenly distributed and U: pike initially not evenly distributed.

The steady-state pike distribution was analyzed based on three different approaches: the pike distribution over the different HSI values in the study area, the pike distribution across the study area and the Shannon-Wiener entropy of the pike distributions across the study area, H ([Shannon, 1948](#_bookmark27)), which calculates the entropy as follows:

expected number of pike in the spawning grounds due to a very low rel- ative occurrence in the 10 m grid.

1. Results
   1. *Steady-state pike distribution*
      1. *Distribution across the study area*

The results showed a strong impact of the grid resolution and the pike density on the ﬁnal pike distribution, whereas a weaker effect of the initial pike distribution was observed ([Fig. 4](#_bookmark9)). Speciﬁcally, the CHE and CHU simulations led to similar pike distributions and showed close agreement with the expected distribution over the study area ([Figs. 3](#_bookmark8)A, [4, 9](#_bookmark9)A and [10](#_bookmark15)A) in contrast to other simulations ([Figs. 3](#_bookmark8)A, [4](#_bookmark9) [and 5–8](#_bookmark9)A). For the L simulations, migration did not occur when the ini- tial distribution was even (CLE and FLE) and was restricted to the neighbourhood of the initially occupied cell in simulations where the initial distribution was uneven (CLU and FLU; [Figs. 5 and 6](#_bookmark10)A respective- ly). Although the steady-state pike distribution in the CME simulations did not reﬂect the expected pike distribution, pike already preferred the spawning places and the high-quality habitats in the river over the low- quality habitats. In contrast, in the CMU simulations the high-quality habitats upstream were not occupied ([Fig. 8](#_bookmark13)A). A similar impact of the initial pike distribution was observed in the FM and FH simulations, although in contrast to the CME simulation, the high-quality habitats were not preferred over the low-quality habitats in the FME simulation. The FL and FM simulation results were not presented, because they were similar to the CL and CM simulations.

* + 1. *Entropy*

Only in the CHE, CHU and CMU simulations the ﬁnal entropy differed less than 1.2 from the expected entropy ([Figs. 8, 9 and 10](#_bookmark13)D). The smallest difference (0.2) was observed in the simulation that did not lead to the expected distribution (CMU; [Figs. 4 and 8](#_bookmark9)), emphasizing the need to consider not only the entropy but also the spatial pike distribution to evaluate model performance. Additionally, the difference between the resulting and expected entropy was higher in F simulations, compared to C simulations and lower in U simulations compared to E simulations, except for simulations CHE and CHU, which did not differ ([Figs. 9 and](#_bookmark14) [10](#_bookmark14)). Overall, simulation results showed a smaller inﬂuence of initial pike distribution at higher pike densities ([Figs. 5 and 6](#_bookmark10)).

* + 1. *HSI distribution*

The pike distribution across the HSIs showed close agreement with the expected distribution for simulations with a high pike density, such as the CHE, CHU and FHE simulations, but not in the FHU simula- tion. Although the CM and FM simulations did not lead to an expected pike distribution, their distribution across the HSIs did approximate the expected HSI distribution when pike were initially evenly distribut- ed (CME and FME) ([Fig. 7](#_bookmark12)B).

* 1. *Number of iteration steps before convergence*

The number of iteration steps needed to obtain convergence was

*P ni*

(*ni*\

affected by the number of pike in the grid, the initial pike distribution

*H* ¼ −∑*i*¼1 *N log N*

ð2:4Þ

and the grid resolution. At high grid resolution (F simulations), more iteration steps were needed in the E than U simulations, whereas a

where *P* is the number of cells, *N* is the total number of pike in the grid and *ni* is the number of pike per cell. For every iteration step, *H* can be calculated and compared to the expected *H* ([Figs. 5–12](#_bookmark10)D). The expected distribution was calculated as the product of the total number of pike and the likelihood that a pike individual would end up in a certain cell, which is the product of the HSI class and the frequency of its occur- rences, scaled between 0 and 1. Consequently, pike are not necessarily most likely to end up in the spawning grounds, although these have the highest HSI ([Fig. 3](#_bookmark8)B and D). Accordingly, [Fig. 3](#_bookmark8)B shows a very low

course resolution required less iteration steps in the E than U simulations ([Fig. 4](#_bookmark9)). For those simulations that led to the expected pike distribution and entropy, 1200 (CHE) and 3400 (CHU) iteration steps were needed before convergence was reached ([Figs. 4, 9](#_bookmark9)C and [10](#_bookmark15)C), whereas C simu- lations with a lower pike density (CL and CM) showed faster conver- gence. By comparison, the maximum number of iteration steps needed to reach convergence in the F simulations was only 10,000, although the grid resolution was 50 times higher than in the C simulations. This indicates that convergence was reached in the former, before the

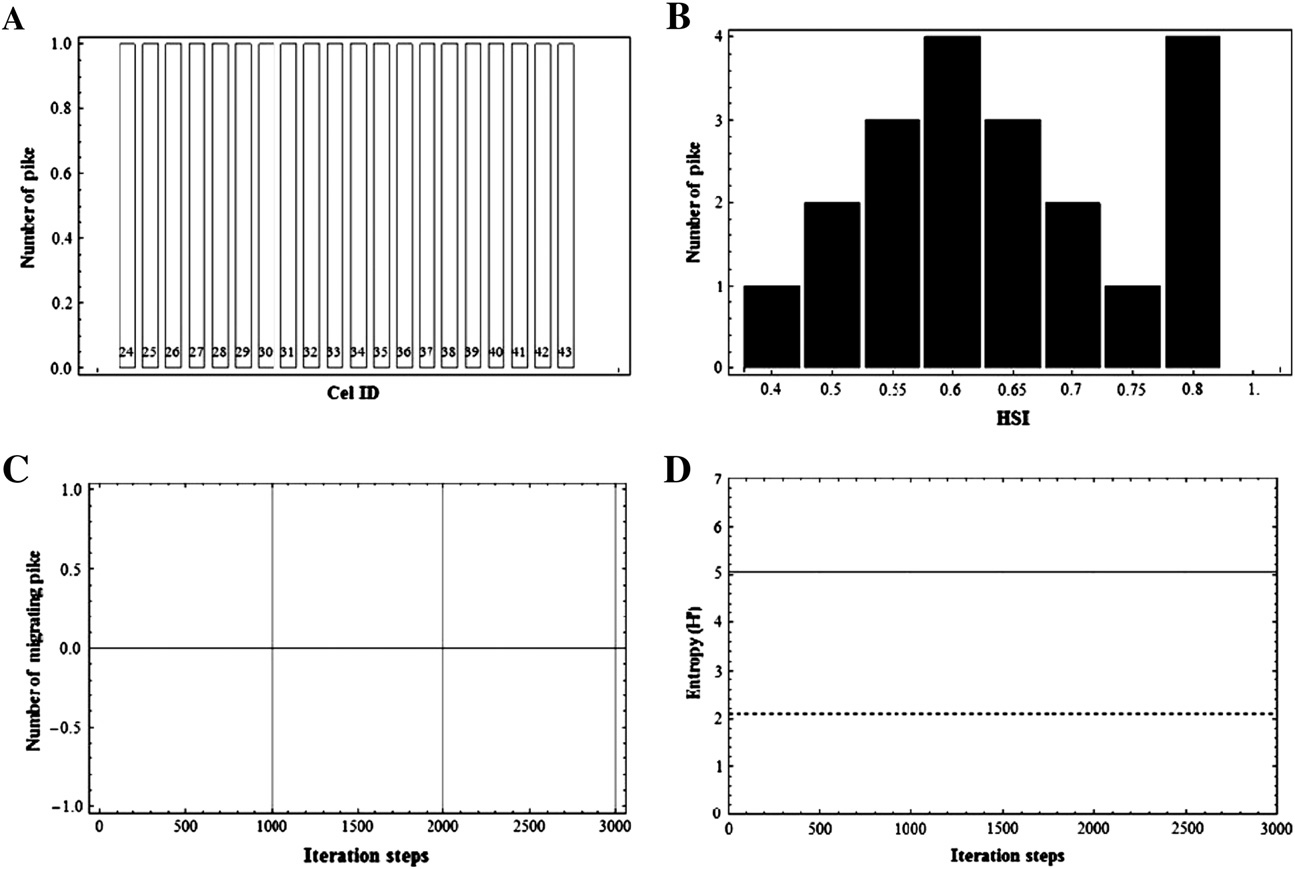


Fig. 5. Output of the CLE simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating be- tween two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected *H*.

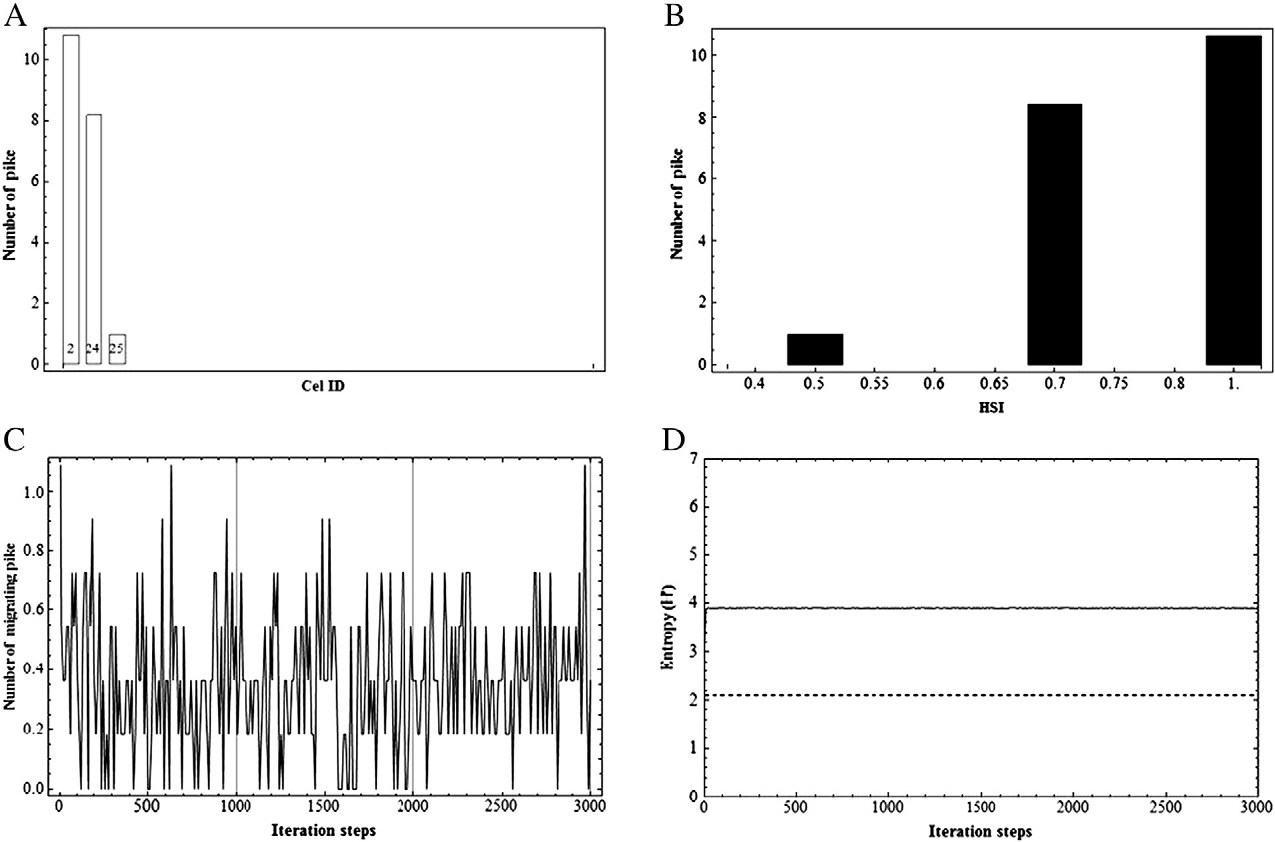


Fig. 6. Output of the CLU simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating be- tween two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy H of the pike distribution over the study area (D). The dashed line indicates the expected H.

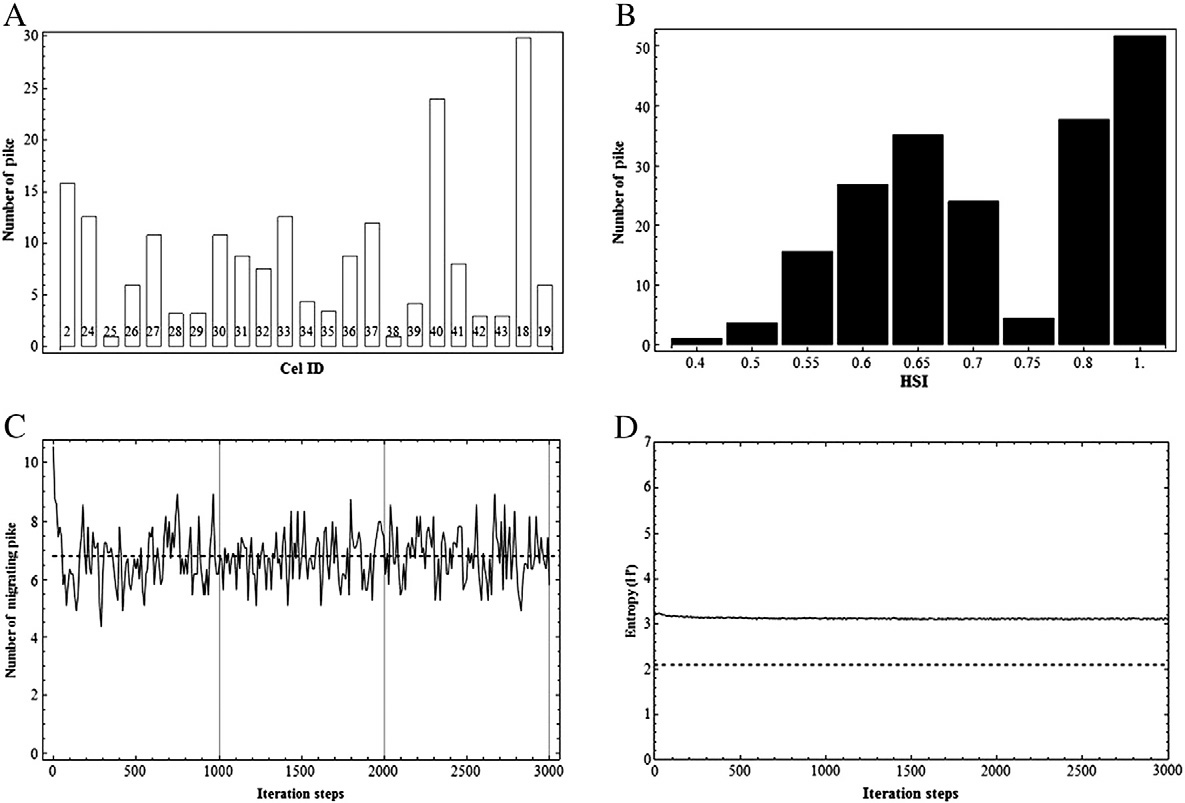


Fig. 7. Output of the CME simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating between two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected *H*.

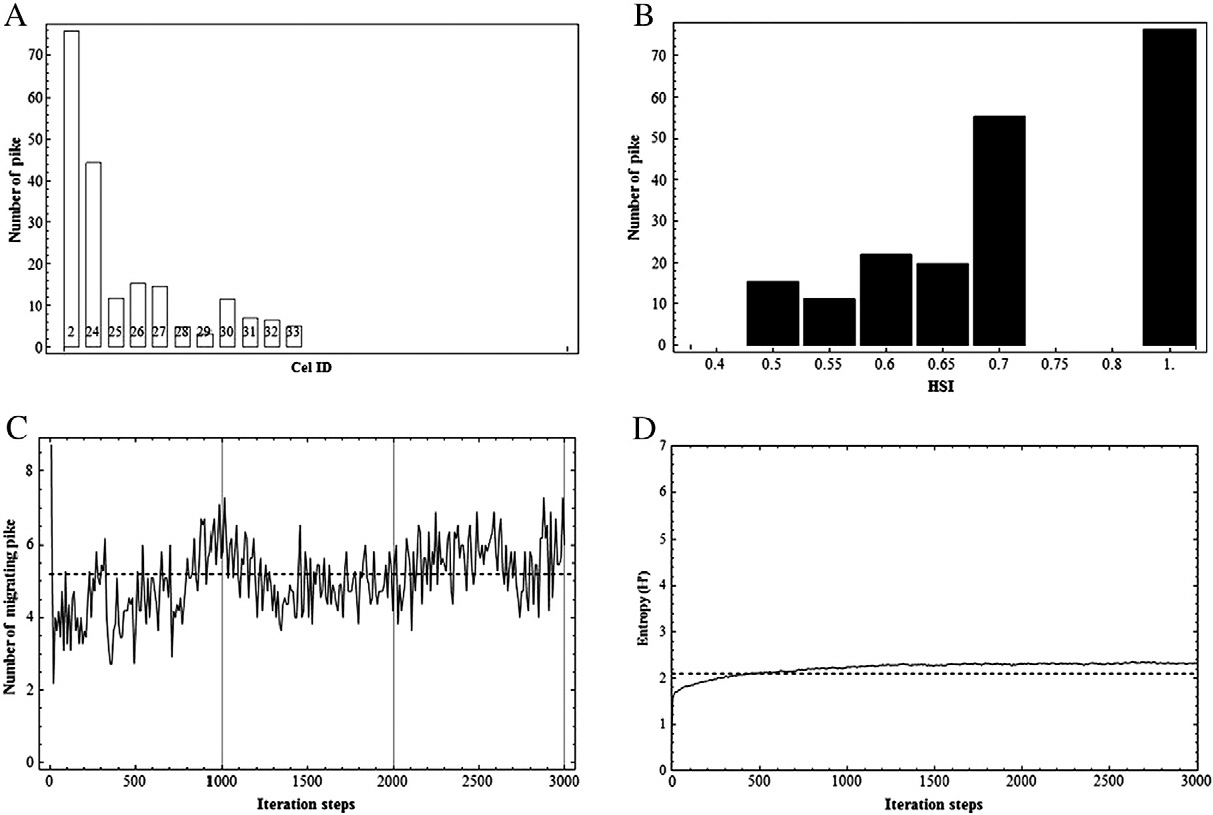


Fig. 8. Output of the CMU simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating between two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected *H*.

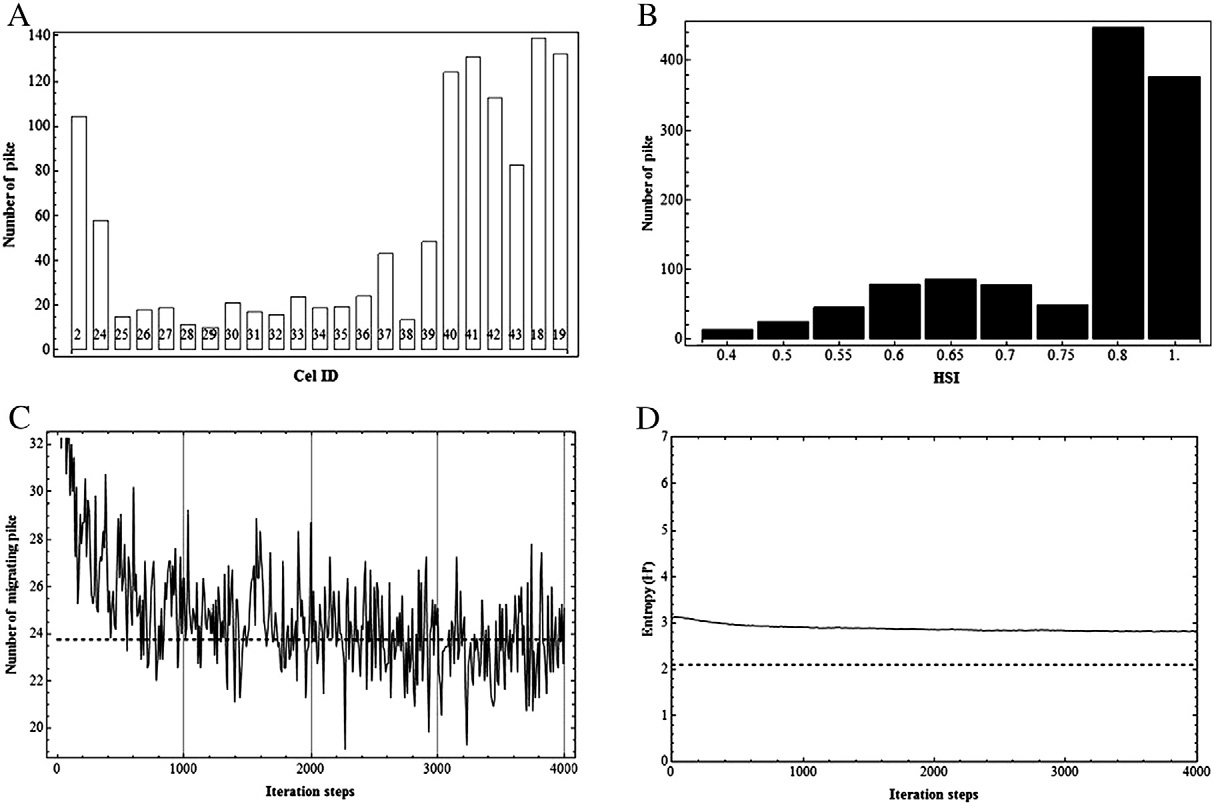


Fig. 9. Output of the CHE simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating between two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected *H*.

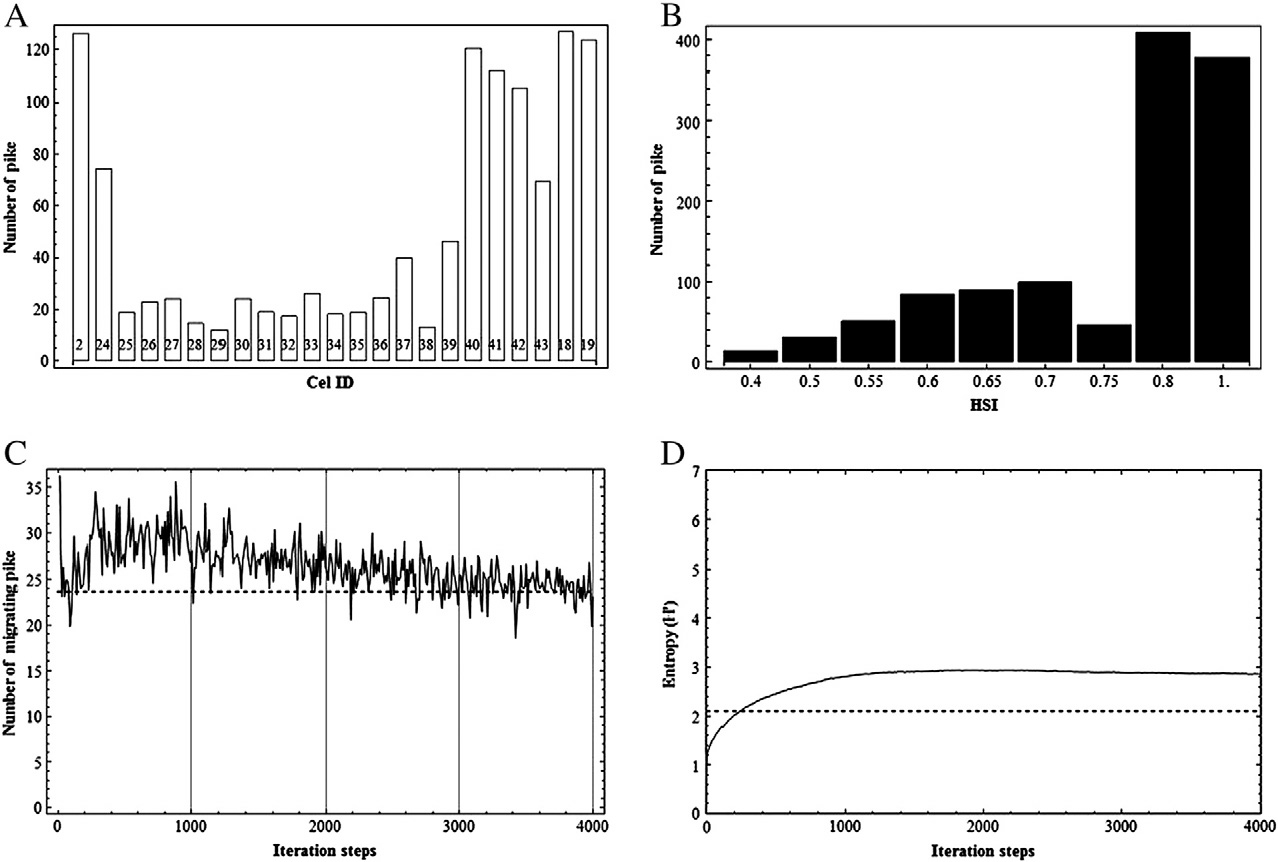


Fig. 10. Output of the CHU simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating between two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected *H*.

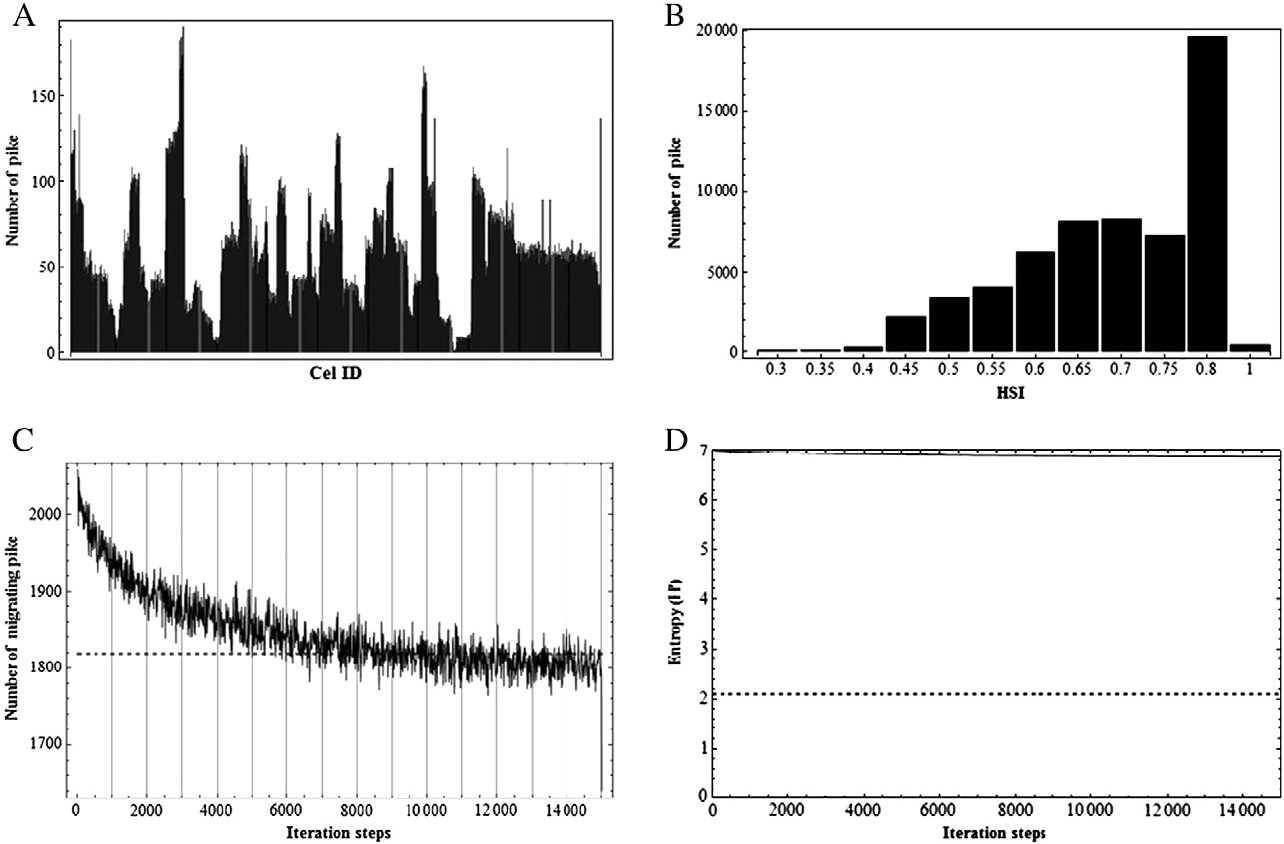


Fig. 11. Output of the FHE simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating between two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected *H*.

expected pike distribution was obtained, which was conﬁrmed by our previous results. For instance, the FME and FLU simulations already reached convergence after 6000 and 2000 iteration steps, respectively, and the FMU and FHU simulations already after 100 iteration steps ([Fig. 4](#_bookmark9)), but none of these simulations led to the expected pike distribu- tion. At low pike density no net migration occurred (simulations CLE, CLU and FLE; [Fig. 4](#_bookmark9)).

1. Discussion
   1. *Simulations of the HSI driven migration*

In 12 simulations we examined the effect of the number of pike in the grid, the initial pike distribution and the grid resolution. The mutually divergent outcomes of these simulations strongly indicate the sensitivity of the model to the number of pike in the grid and to the grid resolution. The results of the CH simulations showed closest agreement with the expected pike distribution. A possible explanation may be that these simulations possess an optimal trade off between the number of pike in the grid and the grid resolution. In other simula- tions, the model may converge because the cell densities decline to one, whereas in the CH simulations, convergence only occurs because the expected pike distribution is reached. Speciﬁcally, the results of the CLE simulation indicate that migration is prevented by densities of one in subsequent cells. Consequently, in this case migration stabilizes before the expected distribution is reached. The likelihood that cell densities equal one is higher at low initial pike densities and high grid resolutions. Additionally, it is also interesting to note that although the pike densities are equal for CH and FH simulations, the latter did not lead to an expected pike distribution. A plausible explanation may be that at a higher resolution, clusters of cells with a similar HSI are

more likely to occur. The transition rules are not adequate to simulate migration through such clusters and lead to a net migration that is zero or completely random. This is an interesting ﬁnding that supports the view of [Chen et al. (2002)](#_bookmark18), who state that it is important, but also difﬁcult to couple the different physical and ecological scales of an eco- system spatially and temporally. Overall, the results clearly illustrate an impact of the model settings on the simulation outcome and conse- quently indicate the importance of choosing the appropriate settings, which is in line with the ﬁndings of [Chen and Mynett (2003)](#_bookmark18) and [Li](#_bookmark18) [(2009)](#_bookmark18). Additionally, the results presented emphasize that although the spatial scales of environmental data are linked properly, model sen- sitivity to model characteristics such as the total number of pike in the grid and the initial pike distribution should be determined before fur- ther analyses are conducted. Moreover, [Graeme (1996)](#_bookmark28) found that also the timing of biological events in a CA can inﬂuence model results. Therefore we believe this should be analyzed in particular before any time-dependent variation is added to the CA.

In the present study, environmental variable selection was only based on expert knowledge. However, when a model is developed to offer recommendations to local managers about a speciﬁc aquatic system, data may provide valuable additional information. Several data-driven techniques, such as classiﬁcation ([Gregori et al., 2010;](#_bookmark30) [Kampichler et al., 2010](#_bookmark30)) and regression trees ([D'heygere et al., 2003,](#_bookmark20) [2006; Pauwels et al., 2010](#_bookmark20)), artiﬁcial neural networks ([Dedecker et al.,](#_bookmark19) [2004, 2005; Lek and Guegan, 1999](#_bookmark19)), support vector machines ([Burges,](#_bookmark18) [1998; Kampichler et al., 2010](#_bookmark18)), logistic regression ([Fransen et al., 2006;](#_bookmark23) [Kampichler et al., 2010](#_bookmark23)) and fuzzy logic ([Adriaenssens et al., 2004,](#_bookmark18) [2006; Mouton, 2008; Mouton et al., 2008, 2009a, 2009b](#_bookmark18)) in combination with genetic algorithms can be applied to select the main inﬂuencing en- vironmental factors. Of course, not only these data mining techniques but also GIS datasets ([Blecic et al., 2009; Li and Gar-On Yeh, 2000;](#_bookmark18)

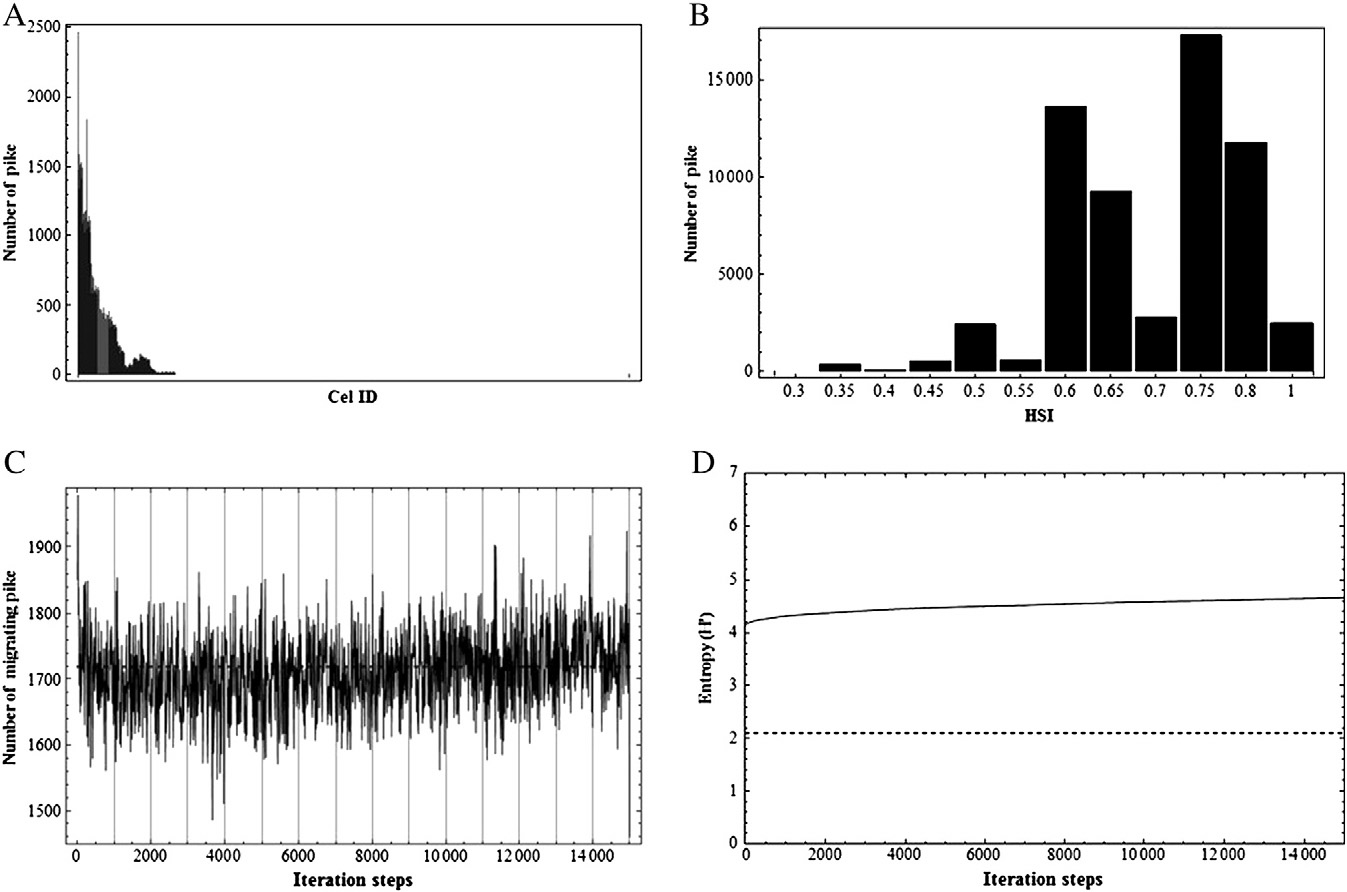


Fig. 12. Output of the FHU simulation showing the pike distribution over the study area (A), the number of pike per Habitat Suitability Index (HSI) value (B), the number of pike migrating between two consecutive time steps (averaged over ten consecutive time steps, C) and the entropy *H* of the pike distribution over the study area (D). The dashed line indicates the expected H.

[Yassemi et al., 2008](#_bookmark18)) could easily be incorporated in a future model thanks to the discrete nature of CAs ([Fig. 2](#_bookmark6)). Consequently, this approach could allow for the integration of more environmental variables ([Hogeweg, 1988; Toffoli, 1984](#_bookmark18)).

To further increase the model reliability, the time dependence of the HSI and a population growth module could be included. Such a dynamic HSI could depend on the season and affect pike differently based on its life stage ([Inskip, 1982](#_bookmark18)); whereas a population growth model could integrate the system's carrying capacity and relations between different life stages. Although only one life stage was consid- ered in the current implementation, other life stages, such as larvae and juveniles, should be integrated if a population growth model would be included. To model the population growth, a deterministic Leslie matrix model could be used, which classiﬁes the population into life stages and simulates the abundances of each stage during a predeﬁned number of discrete time steps ([Caswell, 2001](#_bookmark18)). The Leslie matrix is very suitable to model a species that has only one breeding season per year, such as pike ([Raat, 1988](#_bookmark29)), and could easily be incorporated into the current discrete CA model. Including carry- ing capacity could also stimulate further ﬁne-tuning of the initial number of pike in each cell.

For the simulations presented in this paper, the initial number of pike was chosen based on expert knowledge. However, future data collection could lead to a quantiﬁcation of the carrying capacity of the studied system and thereby increase the accuracy of the initial abundance ([Craig, 1996; Margenau et al., 2008](#_bookmark18)). It is hereby impor- tant to note that although the presented results demonstrated the ef- fect of the grid resolution on the model outputs, the resolution should be chosen primarily based on knowledge on the carrying capacity, the

variability of the environmental variables and the scale of the addressed management issues. Constraints on the resolution could have conse- quences for the simulation of migration when temporally varying events are included in the model, because pike can only shift one cell per time step if a classical neighbourhood is employed. Future research is desir- able to assess these consequences.

Our results indicate that the proposed CA model performs best if grid resolution is low and pike density is high. It is thus crucial to set the optimal trade-off between grid resolution and total number of pike in the grid before any further analyses are conducted or before dynamic processes are included. If model settings are carefully chosen, CAs may be practical to simulate pike migration in a river because they allow river managers to account for the effect of migration barriers on ﬁsh, in contrast to species distribution models that lack the explicit link between grid cells. However, depending on the research objectives, other modelling techniques may be more appropriate to simulate pike migration, particularly when simulations at a high resolution are re- quired. Future research could reveal whether adaptation of the CA tran- sition rules or other modelling techniques, such as IBM, may be more suitable to model populations at a higher spatial resolution. Further, future research could focus on the integration of a growth module and the temporal dynamics of the relation between environmental factors and pike distribution.

Acknowledgements

I. Pauwels is funded by the Flemish Institute for the promotion of Scientiﬁc-Technological Research in the Industry (IWT Flanders).

J. Baetens, I. Pauwels and S. Van Nieuland attended the NKS Summer

School 2010 organized by Wolfram Inc. and speciﬁcally want to thank the staff members for the guidance, Todd Rowland in particular.

References

Adriaenssens, V., De Baets, B., Goethals, P.L.M., De Pauw, N., 2004. Fuzzy rule-based models for decision support in ecosystem management. Science of the Total Envi- ronment 319, 1–12.

Adriaenssens, V., Goethals, P.L.M., De Pauw, N., 2006. Fuzzy knowledge-based models for prediction of Asellus and Gammarus in watercourses in Flanders (Belgium). Ecological Modelling 195, 3–10.

Blecic, I., Borruso, A., Cecchini, A., D'Argenio, A., Montagnino, F., Trunﬁo, G.A., 2009. A cellular automata-ready GIS infrastructure for geosimulations and territorial analysis. Proceedings of the International Conference of Computational Science and its Applications, 1.June-July. Kyung Hee University, Suwon, Korea, pp. 313–327.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121–167.

Casselman, J.M., Lewis, C.A., 1996. Habitat requirements for northern pike (*Esox lucius*).

Canadian Journal of Fisheries and Aquatic Sciences 53, 161–174.

Caswell, H., 2001. Matrix population models, second ed. Sinauer Associates, Sunderland, USA.

Chen, Q.W., Mynett, A.E., 2003. Effects of cell size and conﬁguration in cellular automata based prey–predator modelling. Simulation Modelling Practice and Theory 11, 609–625.

Chen, Q., Mynett, A.E., 2006. Hydroinformatics techniques in eco-environmental modelling and management. Journal of Hydroinformatics 8, 297–316.

Chen, Q., Mynett, A.E., Minns, A.W., 2002. Application of cellular automata to modelling competitive growths of two underwater species Chara aspera and Potamogeton pectinatus in Lake Veluwe. Ecological Modelling 147, 253–265.

Chen, Q.W., Han, R., Ye, F., Li, W.F., 2011. Spatio-temporal ecological models. Ecological Informatics 6, 37–43.

Coeck, J., 2002. Het verschijnsel vismigratie in Vlaamse waterlopen. Referaat Studiedag Vismigratie, 14 juni 2000. Institute of Nature and Forest Research-INBO.

Cooper, J.E., Mead, J.V., Farrell, J.M., Werner, R.G., 2008. Potential effects of spawning habitat changes on the segregation of northern pike (*Esox lucius*) and muskellunge (*E. masquinongy*) in the Upper St. Lawrence River. Hydrobiologia 601, 41–53.

Craig, J.F., 1996. Pike: Biology and Exploitation. Chapman & Hall, London, UK. Craig, J.F., 2008. A short review of pike ecology. Hydrobiologia 601, 5–16.

De Laak, G.A.J., van Emmerik, W.A.M., 2006. Kennisdocument snoek, *Esox lucius* (Lin- naeus, 1758). Sportvisserij Nederland, Bilthoven.

De Nayer, B., Belpaire, C., 1997. De visfauna op de IJzer in 1996. Water 97, 291–300. Dedecker, A.P., Goethals, P.L.M., Gabriels, W., De Pauw, N., 2004. Optimization of Arti-

ﬁcial Neural Network (ANN) model design for prediction of macroinvertebrates in the Zwalm river basin (Flanders, Belgium). Ecological Modelling 174, 161–173. Dedecker, A.P., Goethals, P.L.M., D'heygere, T., Gevrey, T., Leks, S., De Pauw, N., 2005. Application of artiﬁcial neural network models to analyze the relationships be- tween *Gammarus pulex* L. (Crustacea, Amphipoda) and river characteristics. Envi-

ronmental Modelling and Software 111, 223–241.

D'heygere, T., Goethals, P.L.M., De Pauw, N., 2003. Use of genetic algorithms to select input variables in decision tree models for the prediction of benthic macroinvertebrates. Ecological Modelling 160, 291–300.

D'heygere, T., P.L.M., G., De Pauw, N., 2006. Genetic algorithms for optimisation of pre- dictive ecosystems models based on decision trees and neural networks. Ecological Modelling 195, 20–29.

Eklöv, P., 1997. Effect of habitat complexity and prey abundance on the spatial and temporal distribution of perch (*Perca* ﬂ*uviatilis*) and pike (*Esox lucius*). Canadian Journal of Fisheries and Aquatic Sciences 54, 1520–1531.

Farrell, J.M., Mead, J.V., Murry, B.A., 2006. Protracted spawning of St Lawrence River northern pike (*Esox lucius*): simulated effects on survival, growth, and reproduc- tion. Ecology of Freshwater Fish 15, 169–179.

Fransen, B.R., Duke, S.D., McWethy, L.G., Walter, J.K., Bilby, R.E., 2006. A logistic regression model for predicting the upstream extent of ﬁsh occurrence based on geographical in- formation systems data. North American Journal of Fisheries Management 26, 960–975. Glasbergen, M.J., 2001. Fish habitat requirements in the Dutch part of the River Rhine.

WL/Delft Hydraulics Project Report, Delft.

Goethals, P.L.M., Zarkami, R., D'heygere, T., De Pauw, N., Jaarsma, N., Klinge, M., Grimm, M., Van Thuyne, G., Belpaire, C., 2006. Status, ecologie en bescherming van de snoek (*Esox lucius*) in Vlaanderen. Ministerie van de Vlaamse Gemeenschap. Eindrapport TWOL Project nr. B1G/14/2001, studie in opdracht van AMINAL. Afdeling Bos&Groen, Universiteit Gent.

Graeme, D.R., 1996. Effects of the spatial and temporal ordering of events on the be- haviour of a simple cellular automaton. Ecological Modelling 84, 311–314.

Gregori, D., Berchialla, P., Foltran, F., 2010. Comparing bayesian network, artiﬁcial neu- ral networks, classiﬁcation trees and classical logistic models in quantitative risk assessment: an application to the European Registry of Foreign Body Injuries in children. Injury Prevention: Journal of the International Society for Child and Ado- lescent Injury Prevention 16, A216.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology.

Ecological Modelling 135, 147–186.

Haugen, T.O., Winﬁeld, I.J., Vollestad, A.L., Fletcher, J.M., James, J.B., Stenseth, N.C., 2006. The ideal free pike: 50 years of ﬁtness-maximizing dispersal in Windermere. Pro- ceedings of the Royal Society B: Biological Sciences 273, 2917–2924.

Heylen, J., 1997. De hydrologie van het IJzerbekken. Studiedag 'Naar een integraal waterbeleid in het IJzerbekken. Water 97, 239–244.

Hogeweg, P., 1988. Cellular automata as a paradigm for ecological modelling. Applied Mathematics and Computation 27, 81–100.

Inskip, P.D., 1982. Habitat Suitability Index Models: Northern Pike. U.S. Dept. Int., Fish Wildl., Serv. FWS/OB-82/10.17.

Jacobsen, L., Skov, C., Koed, A., Berg, S., 2007. Short-term salinity tolerance of northern pike, *Esox lucius*, fry, related to temperature and size. Fisheries Management and Ecology 14, 303–308.

Kampichler, C., Wieland, R., Calme, S., Weissenberger, H., Arriaga-Weiss, S., 2010. Clas- siﬁcation in conservation biology: a comparison of ﬁve machine-learning methods. Ecological Informatics 5, 441–450.

Kerle, F., Zollner, F., Kappus, B., Marx, W., Giesecke, J., 2001. Fish habitats and vegeta- tion modelling in ﬂoodplains with CASIMIR. CFR project report 13. IWS, University of Stuttgart, Stuttgart, Germany.

Knaepkens, G., Bruyndoncx, L., Coeck, J., Eens, M., 2004. Spawning habitat enhance- ment in the European bullhead (*Cottus gobio*), an endangered fresh water ﬁsh in degraded lowland rivers. Biodiversity and Conservation 13, 2443–2452.

Le Fur, J., Simon, P., 2009. A new hypothesis concerning the nature of small pelagic ﬁsh clusters: an individual-based modelling study of Sardinella aurita dynamics off West Africa. Ecological Modelling 220, 1291–1304.

Lek, S., Guegan, J.F., 1999. Artiﬁcial neural networks as a tool in ecological modelling, an introduction. Ecological Modelling 120, 65–73.

Li, H., 2009. Spatial Pattern Dynamics In Aquatic Ecosystem Modelling. University of Delft, Delft.

Li, X., Gar-On Yeh, A., 2000. Modelling sustainable urban development by the integra- tion of constrained cellular automata and GIS. International Journal of Geographi- cal Information Science 14, 131–152.

Li, W.F., Han, R., Chen, Q.W., Qu, S., Cheng, Z.N., 2010. Individual-based modelling of ﬁsh population dynamics in the river downstream under ﬂow regulation. Ecologi- cal Informatics 5, 115–123.

Maeckelberghe, H., 2002. De kwaliteit van het Vlaamse oppervlaktewater en de afval-waterlozingen in 2002. Een toelichting bij de resultaten van de meetnetten van de Vlaamse Milieumaatschappij. : Report of the Flemish Environment Agen- cy (VMM).

Margenau, T.L., AveLallemant, S.P., Giehtbrock, D., Schram, S.T., 2008. Ecology and man- agement of northern pike in Wisconsin. Hydrobiologia 601, 111–123.

Minss, A.W., Mynett, A.E., Chen, Q., van den Boogaard, H.F.P., 2000. A cellular automata ap- proach to ecological modelling. In: Odgaard, A.J. (Ed.), Proceedings of Hydroinformatics Conference 2000, Iowa, USA.

Mouton, A., 2008. A Critical Analysis of Performance Criteria for the Evalution and Op- timization of Fuzzy Models for Species Distribution. PhD thesis. Ghent University, Ghent, Belgium.

Mouton, A., Schneider, M., Armin, P., Holzer, G., Müller, R., Goethals, P.L.M., De Pauw, N., 2008. Optimisation of a fuzzy physical habitat model for spawning European gray- ling (Thymallus thymallus L.) in the Aare river (Thun, Switzerland). Ecological Modelling 215, 122–132.

Mouton, A., De Baets, B., Goethals, P.L.M., 2009a. Knowledge-based versus data-driven fuzzy habitat suitability models for river management. Environmental Modelling and Software 24, 982–993.

Mouton, A., Jowett, I.G., Goethals, P.L.M., De Baets, B., 2009b. Prevalence-adjusted opti- misation of a fuzzy habitat suitability model for aquatic invertebrate and ﬁsh spe- cies in New Zealand. Ecological Informatics 4, 215–225.

Mouton, A., Gelaude, E., Buysse, D., Stevens, M., Van den Neucker, T., Martens, S., Baeyens, R., Jacobs, Y., Coeck, J., 2009c. Evaluatie van NTMB-projecten langs de IJzer uitgevoerd door W&Z. Luik Visfauna. Studie in opdracht van W&Z, Afdeling Bovenschelde. Rapporten van het Instituut voor Natuur en Bosonderzoek 2009 (INBO.R.2009.61).

Mouton, A., Buysse, D., Stevens, M., Van den Neucker, T., Coeck, J., 2011. Evaluation of riparian habitat restoration in a lowland river. River Research and Applications, <http://dx.doi.org/>[10.1002/rra.1500](http://dx.doi.org/10.1002/rra.1500).

Pauwels, I.S., Everaert, G., Goethals, P.L.M., 2010. Integrated river assessment by cou- pling water quality and ecological assessment models. Proceedings of the iEMSs Fifth Biennial Meeting: International Congress on Environmental Modelling and Software, Ottawa, Canada, pp. 1–9.

Perry, G.L.W., Enright, N.J., 2007. Contrasting outcomes of spatially implicit and spatial- ly explicit models of vegetation dynamics in a forest-shrubland mosaic. Ecological Modelling 207, 327–338.

Provoost, T., 1997. De IJzer: een kleine stroom met een groots verleden. Water 97, 233–238.

Raat, A.J.P., 1988. Synopsis of biological data on the northern pike, *Esox lucius* (Linnaeus 1758). Organisator ter verbetering van de binnenvisserij. FAO of the United Na- tions, Rome. FAO Fisheries Synopsis.

Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423.

Smith, B.M., Farrel, J.M., Underwood, H.B., Smith, S.J., 2007. Year–class formation of Upper St. Lawrence River northern pike. North American Journal of Fisheries Man- agement 27, 481–491.

Toffoli, T., 1984. Cellular automata as an alternative to (rather than an approximation of) differential equations in modeling physics. Physica D 10, 117–127.

U.S. Fish and Wildlife Service, 1981. Standards for the development of habitat suitabil- ity index models for use in the Habitat Evaluation Procedures, U.S.D.I. Fish and Wildlife Service. Division of Ecological Services. ESM 103.

Uchmanski, J., Aikman, D., Wyszomirski, T., Grimm, V., 1999. Individual-based model- ling in ecology — preface. Ecological Modelling 115, 109–110.

Vandenabeele, P., Belpaire, C., Denayer, B., Beyens, J., Yseboodt, R., Samsoen, L., Van Thuynen, G., Maes, J., 1998. De visstand en de bevissing van de Vlaamse binnenwateren. Water 100, 14.

Wolfram, S., 1983. Statistical mechanics of cellular automata. Reviews of Modern Phys- ics 55, 601–644.

Yassemi, S., Dragicevic, S., Schmidt, M., 2008. Design and implementation of an inte- grated GIS-based cellular automata model to characterize forest ﬁre behaviour. Ecological Modelling 210, 71–84.

Yemane, D., Shin, Y.J., Field, J.G., 2009. Exploring the effect of marine protected areas on the dynamics of ﬁsh communities in the southern Benguela: an individual-based modelling approach. ICES Journal of Marine Science 66, 378–387.

Zarkami, R., 2008. Habitat Suitability Modelling of Pike (Esox lucius) in Rivers. PhD the- sis. Ghent University, Ghent, Belgium.

.