



HAL
open science

Distributed MDP for water resources planning and management in inland waterways

Guillaume Desquesnes, Guillaume Lozenguez, Arnaud Doniec, Eric Duviella

► **To cite this version:**

Guillaume Desquesnes, Guillaume Lozenguez, Arnaud Doniec, Eric Duviella. Distributed MDP for water resources planning and management in inland waterways. IFAC World Congress 2017, Jul 2017, Toulouse, France. hal-01577234

HAL Id: hal-01577234

<https://hal.science/hal-01577234v1>

Submitted on 25 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed MDP for water resources planning and management in inland waterways

Guillaume Desquesnes^{*,**} Guillaume Lozenguez^{*,**}
Arnaud Doniec^{*,**} Éric Duviella^{*,**}

^{*} Mines Douai IA, F-59508 Douai, FRANCE
(e-mail: firstname.lastname@mines-douai.fr)

^{**} Univ. Lille, F-59000 Lille, FRANCE

Abstract: Inland waterway management should undergo heavy changes due to a commitment to increase the waterway traffic in a context of climate change. These new constraints will impose an adaptive and resilient management of the water resource. The aim is to plan optimally the water resource distribution over the integrity of the inland waterway network, while taking into account the uncertainties arising from their operation. Due to the large size of waterways, a centralized modeling would not be able to represent an entire network. A distributed Markov Decision Process modeling of inland waterways associated with a resolution algorithm is proposed to allow full scalability, at the cost of optimality. The proposed approach is tested on a subnetwork composed of 7 reaches.

Keywords: Markov decision processes, Multi-agents systems, Distributed artificial intelligence, Large scale systems, Inland waterways, Natural resources management.

1. INTRODUCTION

As shown by the agreement of the COP21 signed in Paris the 12th of December 2015, climate change is a major preoccupation in modern society. In the domain of transportation, the last IPCC *Intergovernmental Panel on Climate Change* report advocates new adaptive measures, particularly by promoting alternative transportation modes (IPCC, 2014). Several studies were carried out on the impact of climate change on inland waterway traffic. The general agreement is that the intensity and the occurrence of flood and drought periods will increase (EnviCom, 2008; IWAC, 2009; Pachauri et al., 2014). In parallel, the fluvial traffic is expected to grow by 35% by 2050 (Beuthe et al., 2014).

The design of new inland waterway management strategies taking into account those new constraints is a priority for the inland waterway managers. Those strategies help deciding when, where and how much water has to be displaced in the network to improve its navigation conditions. Inland waterway networks are large scale systems built by humans to responds to their needs. Those have strong interactions with their natural surroundings. Under the hypothesis of a complete knowledge of these interactions, approaches of adaptive management of inland waterway network in a context of climate changes have been proposed in Nouasse et al. (2015, 2016) using constraints satisfaction problems or quadratic optimization techniques (Duviella et al., 2016). However most of those interactions are only partially known. They are illegal rejects, exchange with groundwater tables, influence of local weather phenomena, ... The management of such networks is subject to numerous uncertainties. For these

reasons, an approach based on a stochastic modeling seems more suitable.

A stochastic approach using Markov Decision Processes to model the inland waterway network and optimize the resource allocation of the whole network while taking into account the numerous uncertainties has been proposed in Desquesnes et al. (2016). However, such a centralized modeling suffer from the lack of scalability of the MDP framework and is unable to scale to real life application. To bypass those limitations a multi-agent modeling based on MDP and a distributed resolution algorithm are proposed.

In this paper, the problem of inland waterway network management in a context of climate change is presented in section 2. Markov decision processes are presented in section 3. Network modeling and distributed resolution algorithm is proposed in section 4. Finally, the proposed algorithm is tested on a realistic scenarios and the results are presented in section 5.

2. WATERWAY NETWORK MANAGEMENT

An inland waterway network (*see* Figure 1) is a large-scale system, mostly used for navigation. It provides both economic and environmental benefits (Mallidis et al., 2012; Mihic et al., 2011) while providing quiet, efficient and safe transports of goods (Brand et al., 2012).

It is mostly composed of canalized rivers and artificial channels, and is divided by locks. Any part between two locks is called a navigational reach. For the sake of simplicity, a navigational reach will be called reach on the rest of this paper.

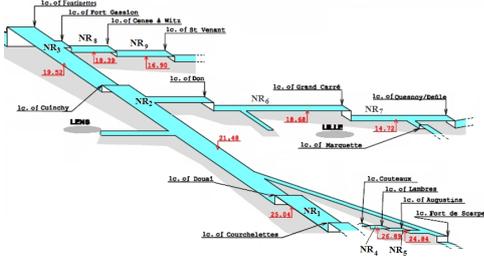


Fig. 1. Part north of France inland waterway network

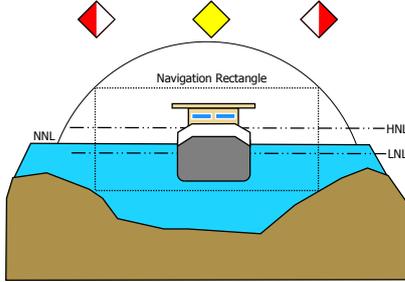


Fig. 2. NNL and navigation rectangle

The main concern of the manager consists in maintaining acceptable water levels in all the reaches of the inland waterway network to allow navigation. This level must respect the conditions defined by the navigation rectangle (see Figure 2) and be as close as possible from the Normal Navigation Level (NNL). The lower and upper boundaries of the rectangle are respectively the Lower and Higher Navigation Level (LNL & HNL) combined with the maximum submerged and emerged boat size.

In normal situations, having a boat crossing a lock is the main disturbance of the water level, since using a lock drains water from the upstream reach towards the downstream reach. Furthermore, the water level is affected by ground exchanges, natural rivers, weather and other unknown factors, like illegal discharges. Locks are not dedicated to control the water level. Structures, such as gates or dams, are used to send water downstream and when available pumps can be used to send water upstream. Those are the main structures used for the distribution of water resources between the reaches of the network.

Navigation is only allowed during the daytime, with few exceptions, notably on Sunday. Reaches management is based on human expertise gathered over time. However, new constraints due to climate change, mostly stronger and more frequent drought and flood periods, and to a commitment to increase the fluvial traffic should impact current management strategies. The main objective is to anticipate the impact of those new constraints by designing adaptive management approaches to ensure in each point of the network and at each instant the navigation requirements. This involves determining a global planning for the water distribution on the whole network by taking into account the uncertainties of climate events and the navigation demand. Planning the distribution of water resources over several time steps allows better anticipation of possible events. The information on the current state of the waterway network is collected in real time through a network of level sensors equipping the reaches.

3. MARKOV DECISION PROCESS

Markov Decision Process (MDP) is a generic framework modeling control possibility of stochastic dynamic system as probabilistic automaton. The framework is well adapted to the waterway network supervision since the state of the network is fully observable (in terms of water volumes) and the control is uncertain due to uncontrolled water transit.

A MDP is defined as a tuple $\langle S, A, T, R \rangle$ with S and A respectively, the state and the action sets that define the system and its control possibilities. T is the transition function defined as $T : S \times A \times S \rightarrow [0, 1]$. $T(s, a, s')$ is the probability to reach the state s' from s by doing action $a \in A$. The reward function R is defined as $R : S \times A \times S \rightarrow \mathbb{R}$, $R(s, a, s')$ gives the reward obtained by attaining s' after executing a from s .

A policy function $\pi : S \rightarrow A$ assigns an action to each system state. Optimally solving a MDP consists in searching an optimal policy π^* that maximizes the expected reward. π^* maximizes the recursive value function of Bellman equation (Bellman, 1957) defined on each state:

$$V_{it}(s) = \max_{a \in A} \left(\sum_{s' \in S} T(s, a, s') \times (R(s, a, s') + V_{it-1}(s')) \right) \quad (1)$$

Based on the algorithm *Value Iteration* (Puterman, 1994), the last V_{it} obtained is then used to generate the optimal policy with equation 2 (V_0 is initialized at 0).

$$\pi^*(s) = \arg \max_{a \in A} \left(\sum_{s' \in S} T(s, a, s') \times (R(s, a, s') + V_{it}(s')) \right) \quad (2)$$

4. DISTRIBUTED APPROACH

The objective is to plan the best course of actions for the entire network over τ time steps, knowing that the conditions of navigation may evolve over time. For example, the weather might become rainy, increasing the water volume of affected reaches; an unexpected increase of the fluvial traffic on some reaches would imply a greater usage of locks and an increase in water transfers.

A time step will represent a period of 12 hours. Using large time steps allows to consider the water level to be uniform on a reach and to smooth the uncertainties on the traffic and on all the temporal variations.

4.1 Network modeling

The distributed approach use the fact that the transfer points (gates, locks, pump, ...) are naturally distributed over the whole network, to divide the management of the water resource between multiple agent. An agent α , that is modeled by a local MDP (MDP_α) will control a subset of the transfer points, denoted TP_α , and observe only the reaches, denoted Re_α , in the subnetwork affected by those transfer points. All transfer points are controlled by exactly one agent, while reaches can be observed by multiple agents. Two agents able to see a same reach are neighbors. N_α represents the set of agents that are neighbors with agent α and TP_{N_α} all the transfer points of the neighbors connected to Re_α .

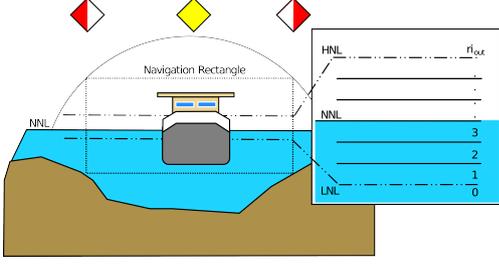


Fig. 3. Discretization of a reach water volume in intervals

A state of MDP_α (agent α) is defined as an assignation of volume for all reaches visible by α , at a given time step. Since the MDP formalism requires a finite number of states, the possible volumes for a reach i are divided in intervals of volumes (see Figure 3). All volumes inferior to the LNL are represented as interval 0 and volumes higher than the HNL correspond to interval ri_{out} . The set of states of agent α can be defined as:

$$S_\alpha = \{0, \dots, \tau\} \times \prod_{i=1}^{|Re_\alpha|} [0, \dots, ri_{out}] \quad (3)$$

The time step τ represents the end of the planning and every state of this last time step is absorbing: $\forall a, T(s, a, s) = 1$, with a null reward: $\forall a, R(s, a, s) = 0$. As time is included into the states, the horizon of resolution can be set to the number of time steps: τ .

Similarly to a state, an action of MDP_α is defined as an assignation of volume transferred by all transfer points affecting Re_α with a part controlled by α and a part controlled by its neighbors N_α . The set of actions of agent α can be defined as:

$$A_\alpha^+ = A_\alpha \times A_{N_\alpha} = \prod_{tp_{i,j} \in TP_\alpha} (tp_{i,j}) \times \prod_{tp_{i,j} \in TP_{N_\alpha}} (tp_{i,j}) \quad (4)$$

where $tp_{i,j}$ represent the set of intervals of the transfer point moving water from reach i to j . Reach 0 represents the unmodelised parts of the network, corresponding to sources supplying the network and destination supplied by the network.

An action $a_\alpha^+ \in A_\alpha^+$ is composed by the action a_α on all controlled transfer points to choose in the combination of possibilities and by a_{N_α} the action given by the neighborhood ($a_\alpha^+ = (a_\alpha, a_{N_\alpha})$). Only a part of the action is controllable by the agent α that induces a coordination problem.

4.2 System dynamic

The aims of an agent is to maintain observed reach within their navigation rectangle while minimizing their distance to their NNL. This is reflected by the local reward function defined on two sub-functions:

$$R_\alpha(s_\alpha, a_\alpha^+, s_\alpha) = R(a_\alpha^+) - \sum_{ri \in Re_\alpha} R(ri_{s'_\alpha})$$

$$R(ri_{s'_\alpha}) = \begin{cases} const & \text{if } ri_{s'_\alpha} \in \{ri_1, ri_{out-1}\} \\ const \times 2 & \text{if } ri_{s'_\alpha} \in \{ri_0, ri_{out}\} \\ NNL_\alpha - ri_{s'_\alpha} & \text{else} \end{cases} \quad (5)$$

where $ri_{s'_\alpha}$ is the volume of reach i during state s'_α . The cost ($NNL_\alpha - ri_{s'_\alpha}$) is replaced by a constant while the interval of volume is partially outside and twice this constant while the interval of volume is outside of the navigation rectangle. It represents the huge penalty of halting the navigation traffic. The function $R : A \rightarrow \mathbb{R}$ models the cost of transferring water from transfer points. This cost function will be specific for each transfer points and each network. They are used to specify some management policies (energy intensive pumping, disturbance of natural water system ...).

Three types of water displacements are specified in the network. Controlled displacements where the volume of displaced water is chosen by the manager at the desired time. Imposed displacements represent the water displacement resulting of the navigation. When a boat want to cross a reach a lock has to be used. Lastly, the uncontrolled displacements correspond to displacements whose quantities are not known by the manager, such as the effect of the weather or illegal discharges.

The uncertainties on the network, imposed and uncontrolled displacements, may vary at each time steps and are therefore represented by temporal variations. A temporal variation v_t is a probability distribution of volumes, displaced by either uncontrolled or imposed displacements, for the whole subnetwork at a time step t . V_t is defined as the list of all possible temporal variations on a time step t and $v_t(i)$ as the probability distribution of volumes affecting reach i for this variation.

The local transition function represents the water displacements and the uncertainties coming from both the discretization of states and actions in intervals and from the temporal variation. Thus it is defined as:

$$LT_\alpha(s_\alpha, a_\alpha^+, s'_\alpha) = \sum_{v_t \in V_t} P(v_t) \times \prod_{i=1}^{|Re_\alpha|} P(ri_{s'_\alpha} | ri_{s_\alpha}, a_i, v_t(i)) \quad (6)$$

with

$$P(ri'_s | ri_s, a_i, v_t(i)) = \begin{cases} p^+ & \text{if } ri_s + a_i + v_t(i) \in ri'_s - 1 \\ p^- & \text{if } ri_s + a_i + v_t(i) \in ri'_s \\ p^- & \text{if } ri_s + a_i + v_t(i) \in ri'_s + 1 \\ 0 & \text{else} \end{cases} \quad (7)$$

where a_i is the resulting of all the actions included in a_α^+ and affecting the reach i .

For a given action a_i on a reach i under specific condition ($ri_s, v_t(i)$), the expected interval is defined as the interval obtained by adding the average values of the tree intervals ri_s, a_i and $v_t(i)$. Due to discretization, reaching the expected interval is not guaranty. p^-, p^+ and p^- correspond respectively to the probability of going to the expected interval, the interval directly superior to the expected and the interval directly inferior to the expected interval, with respect to $p^- + p^- + p^+ = 1$.

Finally the transition function depend on the policies applied by all neighbor agents:

$$T_\alpha(s_\alpha, a_\alpha, s'_\alpha) = \sum_{a_{N_\alpha} \in A_{N_\alpha}} P(a_{N_\alpha} | s_\alpha) \times LT_\alpha(s_\alpha, a_\alpha^+, s'_\alpha) \quad (8)$$

However, the agent α has only a partial view of each neighbor agent β according to the shared observed reaches ($Re_\alpha \cap Re_\beta$). That for the neighborhood action a_{N_α} is approximated as a probability distribution of actions knowing only the local perception of the agent ($P(a_{N_\alpha}|s_\alpha)$).

4.3 Distributed algorithm

The proposed algorithm (see Algorithm 1) is inspired by *LID-JESP* (Nair et al., 2005). It uses agents with local vision whose transitions may depend from the action of other agents. It aims to find iteratively a local joint optimal policy in a distributed way. At each iteration, at most one agent per neighborhood will be able to update its policy. The algorithm will terminate only when all agents are either unable to improve their policies or are in a cycle. To ensure this counters are used. An agent α possess a $counter_\alpha$ initialized to $d > 0$. At the end of an iteration, if an improved policy was available or if the agent is in a cycle the counter will be decreased by one. Else the counter is reset to d (line 14). After that, each agent will share its counter with its neighbors and keep the highest one (line 16). An agent will terminate when its counter reach 0 (line 18).

Algorithm 1 Distributed resolution algorithm

```

1: Create an initial policy  $\pi_0$ 
2:  $it \leftarrow 0$ 
3: repeat
4:   Exchange  $\pi_{it}$  with the neighbors
5:   if Altruistic agents then
6:     Exchange  $R_{it}$  with the neighbors
7:     Build  $R_{it+1}$  with the received rewards function
8:   end if
9:   Build  $T_{it+1}$  from the policies received
10:  Find  $\pi'$  optimal policy of  $M' : \langle S, A, T_{it+1}, R_{it+1} \rangle$ 
11:   $g_{it} \leftarrow gain(\pi', \pi_{it})$  in  $M'$ 
12:   $G_{it} \leftarrow$  neighbors gains
13:   $\pi_{it+1} \leftarrow \pi'$  if  $g_{it} = \max(G_{it})$  else  $\pi_{it}$ 
14:   $counter \leftarrow d$  if  $g_{it} > 0$  else  $counter - 1$ 
15:   $C_{it} \leftarrow$  neighbors counters
16:   $counter \leftarrow \max(C_{it})$ 
17:   $it \leftarrow it + 1$ 
18: until  $counter = 0$ 

```

Proposition. The distributed algorithm will terminate within $d = \max_{\alpha, \beta \in Ag^2} dist(\alpha, \beta)$ iterations iff all agents are in a local optimum or in a cycle. $dist$ is defined as:

$$dist(\alpha, \beta) = \begin{cases} 1 & \text{if } \beta \in N(\alpha) \\ 1 + \min_{\gamma \in N(\alpha)} dist(\gamma, \beta) & \text{else} \end{cases}$$

Proof. Assuming that agent α is not starting iteration c because it finished its resolution at iteration $c - 1$ ($counter_\alpha(c - 1) = 0$), but other agents are not in a local optimum. This implies that iteration $c - d$, there must be at least one agent β who can still improve its policy and thus reset its counter ($counter_\beta(c - d) = d$). Since the counters are decreased by at most 1 at each iteration and are propagated through the neighborhood, at iteration $c - d + dist(\alpha, \beta)$ a lower bound for the counter of agent α can be defined as $c - d + dist(\alpha, \beta)$. Therefore

$counter_\alpha(c - 1) \geq d - dist(\alpha, \beta) + 1 - d + dist(\alpha, \beta) = 1$. So, agent α would need to start iteration c . By contradiction, if the algorithm terminates all agents are in a local optimum or in a cycle.

In the reverse direction, if agents reach a local optimum or a cycle, they would not be able to improve anymore. So the counter is never reset and decrease by 1 at each iteration. After d iterations, all counters are equal to 0 and the agents terminate. \square

At each iteration, every agent will exchange its policy and update its transitions according to the probable neighborhood actions. This new transition function (line 9) will be used to produce a locally optimal policy on the new MDP model. The improvement of this new policy compared to the current policy, in the current model, is evaluated by a heuristic (line 11). In a neighborhood, only the agent with the highest improvement of its policy will be able to keep its new policy (line 13).

For the inland waterway case study, the improvement heuristic $gain_\alpha$ is defined as the difference in the probability to reach the highest rewarded state at each time step using the new and current policies.

$$gain_\alpha(\pi', \pi) = \frac{1}{\tau + 1} \times \sum_{t=0}^{\tau} P(NNL|\pi', t) - P(NNL|\pi, t) \quad (9)$$

The algorithm requires an initial policy for each agent, this case study used a greedy initial policy that chooses the best action for a given state while assuming the absence of neighbors.

This algorithm is defined for egocentric agents, who only aims to optimize the network they manage. Altruistic agents are also defined with an increased perception distance. An altruistic agent α will try to find a policy maximizing both its rewards and its neighbors reward by including in its observed reaches (Re_α) all the reaches impacted by neighbor actions. The reward function of an altruistic agent will be the sum of its local reward function and of approximated reward functions from its neighbors received at the start of each iteration (line 6, 7).

5. RESULTS

A network of 7 reaches and 14 transfer points has been created (see Figure 4). Reaches are represented by squares with their navigation rectangle specified inside. Arcs represent the transfer points and are labeled with their transfer capacity. Every reach is divided in 8 intervals, 6 of size 20 and 2 of infinite size, as specified in section 4.1, while the transfer points are divided in intervals of size 5. The algorithm will plan over 8 time steps. Thus for a single agent approach: $|S| = 8^7 \times (8 + 1)$ and $|A| = 3^6 \times 4^2 \times 6^6$. Therefore a transition function represented by sparse matrix would need to store $|S| \times |A| \times |S| \times \frac{1}{\tau + 1} = 1.81 \times 10^{22}$ values. $\frac{1}{\tau + 1}$ is present, due to the fact that time is linear. A state at time t can only reach a state at time $t + 1$. However, it is not possible to test a non-trivial scenario with a single agent to compare the distributed results to the optimal ones, due to the memory limitations.

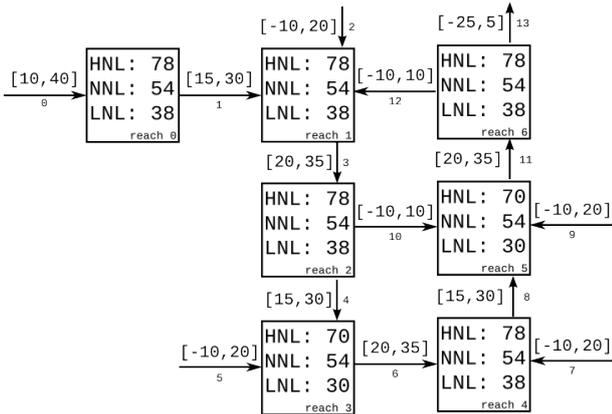


Fig. 4. Seven reaches scenario

decomposition	size	time (s)	out (%)	avg (%)
6 agents	$5.799\,936 \times 10^6$	591	0.000	17.13
7 agents	$2.850\,816 \times 10^6$	167	0.001	15.38
8 agents	$4.751\,36 \times 10^5$	68	0.000	16.45
9 agents	$4.345\,60 \times 10^5$	35	0.000	17.27
10 agents	$3.530\,24 \times 10^5$	32	0.021	17.16
11 agents	$2.919\,68 \times 10^5$	30	0.023	17.83
12 agents	$2.309\,12 \times 10^5$	31	0.016	17.27
13 agents	$1.698\,56 \times 10^5$	30	0.009	16.64
14 agents	$1.088\,00 \times 10^5$	26	0.007	15.57

Table 1. Seven reaches with egocentric agents

decomposition	size	time (s)	out (%)	avg (%)
6 agents	$5.799\,936 \times 10^6$	605	0.000	17.39
7 agents	$2.850\,816 \times 10^6$	446	0.000	18.98
8 agents	$4.751\,36 \times 10^5$	84	0.000	17.86
9 agents	$4.345\,60 \times 10^5$	50	0.000	17.73
10 agents	$3.530\,24 \times 10^5$	33	0.011	18.30
11 agents	$2.919\,68 \times 10^5$	38	0.043	18.91
12 agents	$2.309\,12 \times 10^5$	38	0.021	17.10
13 agents	$1.698\,56 \times 10^5$	35	0.006	17.28
14 agents	$1.088\,00 \times 10^5$	23	0.001	16.28

Table 2. Seven reaches with altruistic agents

The introduced algorithm has been tested with egocentric agents then altruistic agents. For that, multiple decomposition of the network in agents has been proposed by a heuristic algorithm that aims is to minimize the size of the transition function of each agent. To compare the different local joint optimal policies, multiple criteria have been observed. Firstly, the time to obtain the policies and the size of the transition function (sum of all local transition functions). Secondly, two points for the quality of the policy: the percentage of time a reach is outside of its navigation rectangle, *i.e.* out, the average percentage of distance of the water level to the NNL, denoted avg, relatively to the LNL and HLN, at every time step. Finally, the last two criteria are average value of 50000 simulations of the joint policy. This is due to the fact that policies give interval of volume to be displaced by each transfer point. Therefore, in a simulation the value to be transferred by a transfer point is chosen randomly in the interval.

The generation of the policies was made on a cluster using JADE (Bellifemine et al., 1999) framework to develop agents and their protocol.

Results, in Table 1 and 2, show an expected diminution of time and space used to find the local optimal policy when the number of agents increases. In all decomposition, agents reaches almost never get out of their navigation rectangle, while using random volume from the interval of the policy. This leads to the possibility of aberrant values being chosen, putting a reach outside of its navigation rectangle when most values in the interval could have avoided it. Furthermore, the water level of each reach stays relatively close to its NNL, with a percentage of average distance to the NNL strictly inferior to 20%. 20% of distance is close to the size of an action interval and 33% of distance is similar to the size of a state interval.

Figure 5 shows the evolution of the volumes of the reaches in the test network, using the policy generated by the 7 egocentric agents decomposition. The simulation transferred the average value of each interval in each chosen action by the policy. As in Table 1, the average relative distance to the NNL is pretty low (10%). However, it is important to note that it is possible to maintain the NNL at each time step by choosing the correct amount to transfer.

Contrary to what could be expected, increasing the number of agents might not reduce the quality of the local joint optimal policies. When the number of agents increases, agents vision decreases and so from a global point of view, the quality of policies should decrease. However, increasing the number of agents will reduce the uncertainties on the network controlled by each neighbor and the policies they send. This might increase the quality of the policy. Finally, there is no guarantee that for k agents there exists a coherent decomposition of the network.

Egocentric agents (*see* Table 1) seems to produce better results than altruistic agents (*see* Table 2) especially on the relative distance to the NNL. A possible explanation would be that egocentric agents try to maximize a smaller network and so handle less uncertainties than altruistic agents. An egocentric agent will try to maximize only the reaches it affects using approximation on neighbors policies, while an altruistic agent have to maximize the reaches it affects and some reaches visible only to its neighbors using both estimated neighbors policies and estimated states of the unknown reaches. Producing a policy fitting the expected states of every neighbors reaches doesn't seem to be efficient either in computation time or in results quality. However, in all cases the quality of the policies produced is better than the best greedy policy from the initial policies used for those resolutions, with an average relative distance of 37.34% and reaches are outside of their navigation rectangle 6.48% of the time.

In conclusion, results obtained on these scenarios are promising. Reaches stay relatively close to their NNL and shall not go outside of their navigation rectangle when used by a rational user. An egocentric approach seems to yield better results than a more altruistic version, most likely due to the lower number of uncertainties. Alas, no comparison between those results and an optimal result obtained by using one agent was made due to an inability to represent a network of that size.

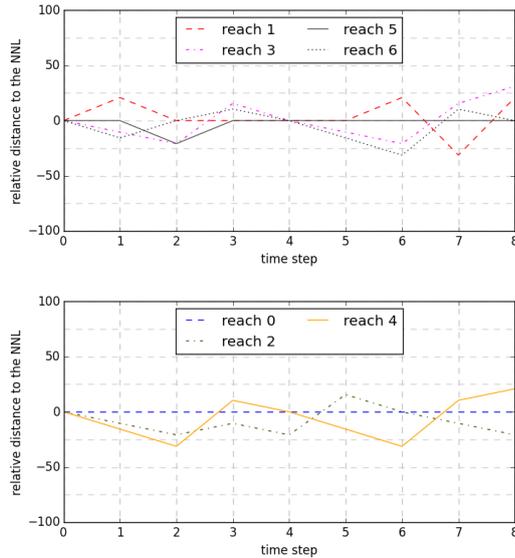


Fig. 5. Evolution of the relative distance to the NNL over the time using the 7 egocentric agents policy

6. CONCLUSION

In this paper, a distributed Markov decision process based approach is presented to optimize the water resource planning and management of inland waterway networks on a given horizon. This approach aims to reduce the impact of drought and flood that may be increased by climate change in the next years.

An agents oriented modeling of the network is introduced to take advantage of the distributed aspect of the inland waterway network. An agent will represent a subset of transfer points and use a MDP to model the dynamic and the uncertainties of the subnetwork affected by those transfer points. There is no guarantee of optimality for the local joint policy obtained by this approach. However, results are encouraging. Increasing the number of agents to model the network, will only guarantee a decrease of computation time and resource usage, up to a certain number of agents where the improvements will stagnate. It is important to note that the quality of the final policy will not necessarily be better for lower number of agents, except for a single agent which gives an optimal policy.

Future works will explore different methods for partitioning the network in agents, with various heuristics to find a good compromise between memory usage, calculation time and deviation from the optimal joint policy. The impact of the initial policy on the final results has to be studied. Furthermore, to reduce the size of the actions intervals, chaining resolution might be considered. By using the joint policy from a first resolution as the new action domain of a new distributed MDP, a smaller discretization of actions could be achieved with similar memory usage. Finally, using this approach on a real network would allow comparison with expert in the management of the inland waterway network.

REFERENCES

Bellifemine, F., Poggi, A., and Rimassa, G. (1999). Jade—a fipa-compliant agent framework. In *Proceedings of*

- PAAM*, volume 99, 33. London.
- Bellman, R. (1957). A Markovian Decision Process. *Journal of Mathematics and Mechanics*, 6(4), 679–684.
- Beuthe, M., Jourquin, B., Urbain, N., Lingemann, I., and Ubbels, B. (2014). Climate change impacts on transport on the Rhine and Danube: A multimodal approach. *Transportation Research Part D: Transport and Environment*, 27, 6 – 11.
- Brand, C., Tran, M., and Anable, J. (2012). The UK transport carbon model: An integrated life cycle approach to explore low carbon futures. *Energy Policy*, 41, 107–124. doi:10.1016/j.enpol.2010.08.019.
- Desquesnes, G., Lozenguez, G., Doniec, A., and Duviella, E. (2016). Dealing with large MDPs, case study of waterway networks supervision. In *Advances in Practical Applications of Scalable Multi-agent Systems. The PAAMS Collection*, 48–59. Springer.
- Duviella, E., Nouasse, H., Doniec, A., and Chuquet, K. (2016). Dynamic optimization approaches for resource allocation planning in inland navigation networks. *ETFA2016, Berlin, Allemagne, September 6-9*.
- EnviCom (2008). Climate change and navigation - waterborne transport, ports and waterways: A review of climate change drivers, impacts, responses and mitigation. *EnviCom - Task Group 3*.
- IPCC (2014). Climate change 2014. *The Core Writing Team, R. K. Pachauri and L. Meyer, Synthesis Report*.
- IWAC (2009). Climate change mitigation and adaptation. implications for inland waterways in england and wales. *Report*.
- Mallidis, I., Dekker, R., and Vlachos, D. (2012). The impact of greening on supply chain design and cost: a case for a developing region. *Journal of Transport Geography*, 22, 118–128. doi:10.1016/j.jtrangeo.2011.12.007.
- Mihic, S., Golusin, M., and Mihajlovic, M. (2011). Policy and promotion of sustainable inland waterway transport in Europe Danube River. *Renewable and Sustainable Energy Reviews*, 15(4), 1801–1809. doi:10.1016/j.rser.2010.11.033.
- Nair, R., Varakantham, P., Tambe, M., and Yokoo, M. (2005). Networked Distributed POMDPs: A Synthesis of Distributed Constraint Optimization and POMDPs. In *National Conference on Artificial Intelligence*, 7.
- Nouasse, H., Doniec, A., Duviella, E., and Chuquet, K. (2016). Efficient management of inland navigation reaches equipped with lift pumps in a climate change context. *4th IAHR Europe Congress, Liege, Belgium 27-29 July*.
- Nouasse, H., Rajaoarisoa, L., Doniec, A., Duviella, E., Chuquet, K., Chiron, P., and Archimede, B. (2015). Study of drought impact on inland navigation systems based on a flow network model. In *Information, Communication and Automation Technologies (ICAT), 2015 XXV International Conference on*, 1–6. IEEE.
- Pachauri, R.K., Allen, M., Barros, V., Broome, J., Cramer, W., Christ, R., Church, J., Clarke, L., Dahe, Q., Dasgupta, P., and others (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.
- Puterman, M.L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc.