

Evolutionary Multiobjective Direct Policy Search (EMODPS), eine heuristische Entscheidungshilfe in der Energie-Wasserwirtschaft

Christoph Libisch-Lehner¹

Kurzfassung:

Talsperren und Wasserspeicher wurden bereits vor mehr als 4500 Jahren errichtet, sind heute Teil der modernen Infrastruktur und werden in der Regel im komplexen sozio-ökonomischen Kontext beschrieben. Sie dienen der Energieerzeugung, der Trinkwasserversorgung, der Bewässerung, dem Hochwasserschutz oder der Niederwasseraufhöhung in Trockenperioden. Ein effektiver und optimierter Betrieb der Anlagen ist von großer Bedeutung, um die Ressource Wasser bestmöglich zu nutzen.

Die Komplexität der Einsatzplanung einzelner Anlagen werden durch den Einfluss externer Größen wie: hydro-meteorologische Rahmenbedingungen, Veränderungen im Strommarkt, die Bedürfnisse mehrerer Nutzer gleichzeitig zu bedienen, oder das Zusammenspiel zwischen wasserwirtschaftlichen Anlagen im Verbundsystem, erhöht.

Der Betrieb von wasserwirtschaftlich Anlagen kann mittels der „Theorie zur optimalen Steuerung“ (Optimal Control Theory) beschrieben werden. Diese Theorie setzt sich mit dem Problem auseinander, Kontrollregeln für ein System zu finden, um dieses im Sinne einer Zielfunktion bestmöglich zu betreiben. Traditionell werden optimale Betriebsregeln mittels direkter Optimierungsmethoden (z.B., Dynamische Programmierung) ermittelt. Für komplexe, „real-world“ Probleme weisen diese Methoden jedoch Limitierungen auf. Dazu zählen der „Fluch der Dimension“, der „Fluch der Modellierung“ und der „Fluch der Mehrzielplanung“. Mit Methoden wie EMODPS können die genannten Probleme überwunden werden. Durch drastische Reduktion von Entscheidungsvariablen und durch die getrennte Betrachtung von wasserwirtschaftlichen System und Optimierungsstrategie (indirekte Optimierung) kann die Systemkomplexität ohne Einschränkung erhalten bleiben.

Keywords: Reinforcement Learning, Reservoir Optimierung, Direct Policy Search, Heuristik

1 Einleitung

Wasserspeicher dienen als Ressource für saubere erneuerbare Energie, Versorgen die Gesellschaft mit Trink- und Nutzwasser, dienen als Erholungsraum und bieten Schutz vor Hochwasser. Die älteste, teilweise noch erhaltene Talsperre liegt in Ägypten. Sie wurde vermutlich um 2500 v. Chr. errichtet und sollte der Bewässerung und dem Hochwasserschutz dienen. In Deutschland wurden die ersten Talsperren im Mittelalter und der vorindustriellen Neuzeit errichtet und dienten vorwiegend dem Erzbergbau. Wasser wurde für den Antrieb von Fördermaschinen und Pochwerken, sowie zur Erzwäsche bereitgestellt. Bei den ersten Talsperren handelte es sich meistens um Dämme (DTK, 2013). Der moderne bzw. neuzeitli-

¹ Pöyry Energy GmbH, Kranichberggasse 4, 1120 Wien, +43 664 828 72 21, christoph.libisch@poyry.com, <http://www.poyry.at/hydro-consulting>

che Talsperrenbau wurde in den letzten Jahrzehnten des 19. Jahrhunderts eingeleitet (Lehner et.al. 2011). In diese Zeit fällt auch die Arbeit von Wenzel Rippl („Der Wasserbauer der Donaumonarchie“, Dozent an der TU-Graz) im Jahre 1883. Seine Arbeit gilt als eine erste Pionierarbeit und setzt sich mit der Dimensionierung von Talsperren zur Trinkwasserversorgung auseinander. Damals wie heute waren die Dimensionierung und der Betrieb der Anlagen im Fokus. Zu Beginn war der Betrieb von Wasserspeichern eine Abfolge von Wasservolumina die innerhalb eines Zeitintervalls aufgrund von Referenzbedingungen abgegeben wurden. Später wurde dieser Ansatz durch Regelkurven erweitert, die einen Zielspeicherstand zu jedem Zeitpunkt des Jahres vorgaben. Ein Umdenken passierte in den 1960er Jahren, als unter dem Einfluss der Theorie der optimalen Steuerung und Operational Research, erkannt wurde, dass der Betrieb eines Reservoirs aufgrund von Zustandsgrößen (Speicherzustand) erfolgen soll. Einer der erfolgreichsten Ansätze für die Lösung des Problems ist Dynamische Programmierung (Bellman, 1957). Die erste Implementierung der deterministischen Version erfolgte 1961 durch Hall und Buras. In späterer Folge, aufgrund von verbesserter Computerleistung fand die stochastische Variante von dynamischer Programmierung weite Verbreitung (Zusammenfassung in Castelletti, 2007).

Aufgrund der langen Tradition von Bau und Betrieb wasserwirtschaftlicher Anlagen hat die Forschung auf diesem Gebiet eine ebenso lange Geschichte. Eine Vielzahl von Publikationen wurden in diesem Forschungszweig publiziert – laut Google-Scholar über 500.000 Einträge. Aufgrund der Wichtigkeit und zahlreicher Veränderungen im Wassersektor ist weiterer Forschungsbedarf vorhanden. Als allgemeine Herausforderung ist zunächst der Klimawandel zu nennen. Die Unsicherheit im zukünftigen Abflussgeschehen steigt – mehr intensive und dafür aber auch variabelere Extremereignisse sind zu erwarten. Diese Schwankungen müssen durch Reservoirs ausgeglichen werden. Durch ausgedehnte Trockenperioden ändert sich auch der Wasserbedarf. Ein weiterer Unsicherheitsfaktor ist der sozio-ökonomische Wandel, insbesondere die Veränderungen am Strommarkt sind spürbar. Durch die Einspeisung volatiler Energieträger ändern sich die Preisstruktur und der Energiebedarf. Mehr Flexibilität in der Wasserkrafterzeugung ist gefordert.

Der Neubau von Anlagen, insbesondere im Wasserkraftbereich geht zurück. Der Fokus liegt auf dem Betrieb und der Erhaltung von Wasserkraftwerken. Anforderungen an die Anlagen in punkto Naturschutz bzw. ökologisch verträglichen Betrieb (Stichwort Schwall und Sunk) steigen. Kommt es zu Neubauten werden diese mitunter in ein bestehendes wasserwirtschaftliches System eingegliedert. Die Planung neuer Anlagen und die Bewertung im Sinne eines integrierten Wasserressourcenmanagement erfordert die ganzheitliche Betrachtung von wasserwirtschaftlichen Mehrzweckanlagen. Die gesteigerte Komplexität im Umfeld dieser Anlagen stellt traditionelle Optimierungsverfahren vor Herausforderungen (siehe Kapitel 3.1)

Diesen Herausforderungen stehen neue Möglichkeiten gegenüber. Dazu zählt die höhere Verfügbarkeit von Daten die den Wasserkreislauf beschreiben. In den vergangenen Jahrzehnten wurde das Potential von Messsystemen zur Erdbeobachtung erkannt und ausgebaut. Dies führte zu einem besseren Systemverständnis und verbesserte unter anderem die Prognosemöglichkeit von Zuflüssen. Die erhöhte Rechenleistungen und der Forschungserfolg auf dem Gebiet von evolutionsbasierten Algorithmen, die einfache Einbindung exogener Informationen sowie die klare Erhaltung der Systemkomplexität lassen den Fokus auf simulationsbasierte Optimierungsverfahren fallen.

In der Wasserwirtschaft wurde simulationsbasierte Optimierung mittels genetischer Algorithmen bereits 1997 erprobt (Oliveira und Loucks, 1997) und 2003 als Parameter-Simulation-Optimization beschrieben (Koutsoyiannis und Economou, 2003). Giuliani et. al., 2014 erweiterten den Ansatz um die nichtlineare Parametrisierung der Betriebsstrategie und koppelten Approximationsnetzwerk (neuronaales Netz) an einen selbstanpassenden Evolutionsalgorithmus (Borg MOEA).

2 Problemformulierung

Beim Betrieb eines wasserwirtschaftlichen Systems geht es darum zu wissen wann, wieviel Wasser aus einem Reservoir abzugeben ist. Dabei handelt es sich um eine sequentielle Entscheidungsfindung (Puterman, 2008) unter Unsicherheit da zukünftige Zuflüsse und Nutzungsanforderungen unbekannt sind.

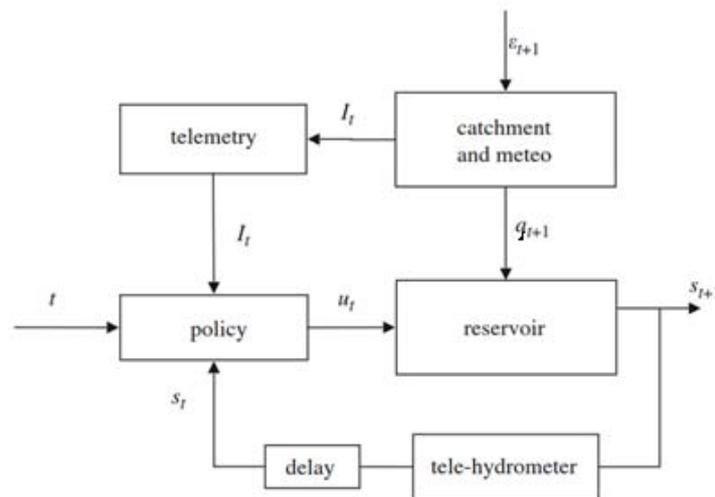


Abbildung 1 Feedback-Control-Framework, Castelletti et. al. 2007

Ein wasserwirtschaftliches System kann durch ein Feedback-Control-Framework (Doyle et. al. 1990, vgl. Abbildung 1) abgebildet werden. Aus einem Systemzustand heraus werden Aktionen/Entscheidungen getroffen. Die Entscheidung hat zur Folge, dass ein Nutzen (oder im Falle einer Fehlentscheidung Kosten) entsteht und sich das System gemäß einer, durch die Aktion bestimmten, Wahrscheinlichkeitsverteilung zu einem neuen Zustand entwickelt. Dies kann durch einen partiell beobachtbaren Markov-Entscheidungsprozess, bei dem alle Entscheidungen, Belohnungen und Systemübergangswahrscheinlichkeiten vom momentanen Systemzustand abhängen, beschrieben werden:

$$x_{t+1} = f_t(x_t, u_t, \varepsilon_{t+1})$$

wobei sich der Zustandsvektor x_t aus den Zustandsvariablen S_t und I_t zusammensetzt. S_t ist der aktuelle Speicherinhalt und I_t beinhaltet z.B. Information über den hydrologischen Sys-

temzustand des Einzugsgebiets. Der hydrologische Systemzustand wird durch zukünftige Zuflüsse oder Daten, die auf zukünftige Zuflüsse schließen lassen (Grundwasserstände, Zustand der Atmosphäre, Schnee-Wasser-Wert, etc.) beschrieben. Alle nötigen Informationen um eine Entscheidung u_t zu treffen werden aus dem Zustandsvektor gewonnen. Es sei erwähnt, dass die Entscheidung u_t nur auf die Veränderung des Speicherzustandes einen Einfluss hat. Aufgrund der Tatsache, dass zukünftige Zuflüsse nur bedingt prognostizierbar sind unterliegt die Information stochastischen Störungen (Unsicherheit, ε_{t+1}). Dieses Problem wird in der Literatur als „hidden state“-Problem bezeichnet (Hasinoff, 2003).

Die Dynamik des Reservoirs wird über die Massenbilanzformel abgebildet:

$$S_{t+1} = S_t + q_{t+1} - e_{t+1} - r_{t+1}$$

wobei S den Speicherinhalt, q den gesamten Zufluss zum Reservoir, e Verdunstungsverluste und r das abgegebene Volumen beschreibt. Der Zeitindex der Variablen beschreibt den Zeitpunkt, wann der Wert deterministisch bekannt ist – S_t wird zum Zeitpunkt t gemessen, während die Variablen q , e und r über den Zeitintervall $[t, t + 1)$ integriert werden. r_{t+1} ergibt sich aus einer Entscheidung u_t , wobei darauf geachtet werden muss dass die tatsächlich Abgabe r_{t+1} von der Entscheidung aufgrund physikalischer Rahmenbedingungen abweichen kann (z.B. wenn das Reservoir leer wird). Die Entscheidung u_t wird aufgrund einer Betriebsstrategie p getroffen. p ist als Sequenz von Kontrollregeln $\{m_0 \dots m_{T-1}\}$ definiert:

$$p = \{m_0, m_1, \dots, m_{T-1}\}$$

$$u_t = m_t(x_t)$$

Im Zuge der Systemevolution (Massenbilanzformel), also der Transformation von x_t zu x_{t+1} entstehen entweder Nutzen (z.B. Erträge aus der Energieerzeugung) oder Kosten (z.B. wenn das Managementziel des Speichers nicht erreicht wird). Diese Kosten/Nutzen werden als gewichtete Summe von k -Zielen ermittelt:

$$g_t = \sum_{j=1}^k w^j g_t^j$$

Die Gesamtkosten/der Gesamtnutzen bezogen auf eine Strategie p und einen Anfangszustand x_0 errechnet sich über den Planungshorizont h wie folgt:

$$J(p, x_0) = \lim_{h \rightarrow \infty} \sum_{t=0}^h \underbrace{E}_{\varepsilon_1 \dots \varepsilon_{t+1}} [g_t(x_t, u_t, \varepsilon_t)]$$

Der langfristige optimale Betrieb eines wasserwirtschaftlichen Systems kann als stochastisches, optimales Steuerungsproblem angesehen werden. Wird der k -dimensionale Zielfunktionsvektor minimiert oder maximiert folgt, dass sich die optimale Strategie p^* aus der Lösung des optimalen Steuerungsproblem ergibt.

$$p^* = \arg \min_p J(p, x_0)$$

3 Lösungswege

Markov-Entscheidungsprozesse können mit „Reinforcement Learning“ Methoden (RL) gelöst werden. Bei RL erlernt ein Agent (z.B. Strategie), durch Interaktion mit seiner dynamischen Umgebung, selbständig eine optimale Strategie. Entscheidungen des Agenten basieren auf trial-and-error, wobei die getroffenen Entscheidungen je nach Auswirkung belohnt oder bestraft werden. RL kann in Methoden unterteilt werden die entweder den Raum der Nutzenfunktion (space of value function) oder den Raum der Betriebsregeln (space of policies) absuchen (Kaelbling, 1996). Die erste Gruppe beinhaltet Methoden der Dynamischen Programmierung. Traditionell werden diese Methoden häufig verwendet um optimale Steuerungsprobleme in der Wasserwirtschaft zu lösen. Die zweite Gruppe nennt man direct policy search (DPS) Methoden. Eine Form von DPS ist EMODPS und wird in Kapitel 3.2 näher beschrieben. DP ist mit einigen Problemen konfrontiert, die im Folgenden erläutert werden. Grundsätzlich scheint, dass DPS derartige Probleme umgehen kann (Kapitel 3.3), jedoch anderen Herausforderungen gegenübersteht (siehe Kapitel 3.4).

3.1 Flüche der Dynamischen Programmierung

In der Praxis wird die Verwendung von DP aufgrund der folgenden „Flüche“ eingeschränkt:

- **Fluch der Dimension:** der Rechenaufwand wächst exponentiell mit der Dimension des Systemzustandes, des Steuerungsvektors und des stochastischen Störvektors (Bellman, 1957).
- **Fluch der Modellierung:** jede Variable, die als Input der Betriebsregel berücksichtigt wird, muss von einem dynamischen Modell beschrieben werden (Bertsekas und Tsitsiklis).
- **Fluch der Mehrzielplanung:** der Rechenaufwand wächst ebenfalls exponentiell mit der Dimension des Zielfunktionsvektors (Powell, 2007).

Zusammengefasst: alle drei Flüche erzeugen einen exponentiellen Anstieg in der Berechnungszeit, um eine Lösung für die rekursive Bellman-Gleichung, die der Kern des DP-Algorithmus ist, zu finden.

DP wird im Zuge eines integrierten Management von Wasserressourcen mit einer neuen Klasse von Problemen konfrontiert. Integriertes Wasserressourcen-Management beschreibt

eine ganzheitliche Bewertung, Planung und Gegenüberstellung von Wasserverfügbarkeit und Wassernutzung. Eines der Hauptziele ist es, die Wasserressourcen effizient zu verwalten und dabei die Interessen aller Beteiligten zu berücksichtigen: sozioökonomischen Einheiten (Wirtschaftssektoren, lokale Bevölkerung, Ressourcenmanager), Verwaltungseinheiten (Länder, Staaten, Provinzen, Bezirke) sowie hydrologische und ökologische Einheiten (Einzugsgebiete und Ökosysteme). Die Einbeziehung unterschiedlicher Interessengruppen erhöht die Dimension der Zielfunktion, die dem Steuerungsproblem zu Grunde liegt. Zudem hat jeder Nutzer wiederum einzelne mehrere Interessen und Vorgaben die erfüllt werden müssen (vgl. Kapitel 4). Zudem ist es nicht möglich, eine optimale Managementpolitik zu entwerfen, indem die Optimierung einzelner Reservoirs separat in Betracht gezogen wird. Bereits für wasserwirtschaftliche Systeme mit drei Reservoirs stellt dies ein Problem dar. Nur durch eine Vereinfachung des Systems können komplexe wasserwirtschaftliche Fragestellungen mittels DP-Methoden gelöst werden. Um jedoch die Systemkomplexität erhalten zu können müssen andere Werkzeuge und Methoden entwickelt und weiterentwickelt werden.

3.2 Evolutionary Multiobjective Direct Policy Search – EMODPS

Evolutionary Multiobjective Direct Policy Search (EMODPS, Giuliani et al. 2014) ist ein simulationsbasierter Optimierungsansatz bei dem DPS, nicht-lineare Approximationsnetzwerke und Mehrziel-Evolutionsalgorithmen für die Erstellung Pareto-optimaler Betriebsstrategien von wasserwirtschaftlichen Systemen kombiniert werden (Abbildung 2).

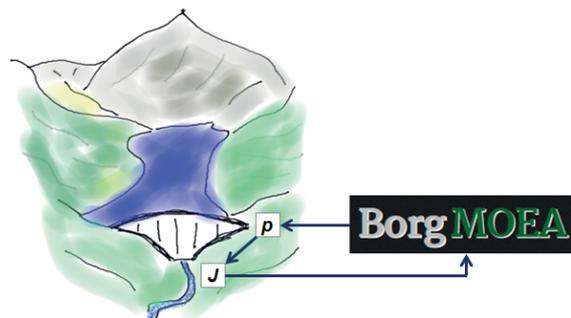


Abbildung 2 Zusammensetzung von EMODPS: das wasserwirtschaftliche System bleibt in seiner Komplexität erhalten; die Betriebsstrategie (p) wird mittels einer nichtlinearen Funktion approximiert; Funktions-Parameter werden mittels BorgMOEA unter Rücksichtnahme des Zielfunktionsvektor (J) optimiert.

Die Betriebsstrategie einer wasserwirtschaftlichen Anlage wird parametrisiert und mit einer geeigneten Funktion abgebildet (Direct Policy Search, Rosenstein und Barto, 2001). Flexible und nichtlineare Funktionen müssen angewendet werden, um zu verhindern, dass sich die Suche nach der optimalen Strategie auf einen Teil des Entscheidungsraums beschränkt, der möglicherweise nicht die optimale Lösung enthält. Untersuchungen (Giuliani et al. 2016) haben gezeigt, dass künstliche Neuronale Netze (kNN), deren Aktivierungsfunktion Radiale Basis Funktionen sind, die besten Ergebnisse liefern. Die Lösung des Optimierungsproblems beschränkt sich auf die Suche nach der optimalen Parametrisierung des kNN's und erfolgt mittels einer Evolutionsstrategie (Moriarty et al., 1999). Bei Evolutionsstrategien handelt es sich um stochastische Suchmethoden, die, in Anlehnung an die biologische Evolution, einen

intelligenten Suchprozess nutzen. Für eine Trainingsperiode (Zeitraum beobachteter oder simulierter Daten (z.B. Zuflüsse oder Strompreise) werden durch evolutionäre Prozesse neue Lösungsstrukturen erschlossen und mittels einer Fitnessfunktion bewertet. Gute Lösungen werden durch die Prozesse Selektion und Vererbung an die nächste Generation von Parametern übergeben. Durch die Vermischung mit einer Zufallskomponente (Mutation) werden die Parameter verändert und neu bewertet. Abhängig von der Komplexität des Optimierungsproblems (Anzahl der Reservoirs, Anzahl von Nutzern, Mehrzielplanung, etc.) müssen geeignete Evolutionsstrategien ausgewählt werden, um eine optimale Lösung zu finden. Zatarain Salazar et al., 2016 haben gezeigt, dass für die Erstellung von Pareto-optimalen Lösungen auf moderne Evolutionsalgorithmen (z.B. Borg MOEA) zurückgegriffen werden sollte. Neben der Parametrisierung der Betriebsregeln und der Kopplung einer geeigneten Evolutionsstrategie ist die Auswahl der Zustandsvariablen, auf deren Basis eine Entscheidung (z.B. Abgabe von Wasser) getroffen wird, von Bedeutung. Diesbezüglich werden Variablen, die den Systemzustand (z.B. Speicherinhalt) beschreiben, verwendet. Zusätzlich kann das kNN mit exogenen Informationen (z.B. Zuflussprognosen, Zeit, Bedarf, Klimavariablen, Strompreisen, etc), die auf zukünftige Systemzustände schließen lassen, betrieben werden. Dadurch wird der Suchraum vergrößert und die Wahrscheinlichkeit die optimale Lösung zu finden steigt.

3.3 Vorteile von EMODPS

Aufgrund der simulationsbasierten Natur hat EMODPS einige Vorteile gegenüber Methoden der DP-Familie. Der größte Vorteil ist, dass die Systemkomplexität (Anzahl der Reservoirs, Zielfunktion) ohne Einschränkung erhalten bleiben kann (vgl. „Fluch der Dimension“). Dies ist insbesondere auf die geringere Anzahl der Entscheidungsvariablen zurückzuführen. Im Fall von EMODPS hängt die Anzahl der Entscheidungsvariablen nur von der Komplexität der parametrisierten Betriebsregel ab. Die Entscheidungsvariablen sind die Parameter der Approximationsfunktion. Weiters können exogene Informationen (Messdaten, die den Systemzustand beschreiben) unkompliziert in die Strategie implementiert werden. Obwohl dadurch der Systemzustand erweitert wird, steigt die Anzahl an Entscheidungsvariablen nicht massiv an (vgl. „Fluch der Modellierung“). Die Rechenzeit steigt nur linear mit der Länge der simulierten Zeitreihe und der Anzahl der Evolutionsläufe. Zudem muss der Wertebereich Entscheidungsvariablen nicht diskretisiert werden (vgl. „Fluch der Dimension“).

Die Komplexität von DP-Methoden, insbesondere ihrer stochastischen Variante, ist die Notwendigkeit der expliziten Kenntnis von Wahrscheinlichkeitsfunktionen der Zustandstransformation (vgl. „Fluch der Dimension“). DPS-Verfahren erfordern diese Funktionen nicht, da die Strategie Kenntnis von der stochastischen Umgebung durch Lernen erwirbt, welches durch die Evolutionsstrategie erzwungen wird (Kaelbling, 1996).

Ein weiterer Vorteil, wenn es sich um Mehrziel-Optimierungsprobleme handelt, ist die Verwendung von multi-objektiven evolutionäre Algorithmen (MOEAs), die eine Schätzung der Pareto-Front in einem einzigen Lauf des Algorithmus ermöglichen. Der Rechenaufwand wächst nicht exponentiell wie bei multi-objektiven DP-Modellen (vgl. „Fluch der Mehrzielplanung“) sondern linear mit der Anzahl der untersuchten Pareto-Optimalen Lösungen.

Im Zuge einer robusten Optimierung (Beyer et. al., 2007) können robuste Betriebsstrategien für komplexe wasserwirtschaftliche Fragestellungen gefunden werden. Robuste Optimierung

ist ein Teilbereich der Optimierungstheorie, der sich mit Optimierungsproblemen beschäftigt, bei denen ein Maß an Robustheit gegen Unsicherheit gesucht wird. Die Unsicherheit wird anhand eines deterministischen Ensembles von Zuflüssen ausgedrückt, wobei eine Optimierung über das gesamte Ensemble erfolgt. Mittels EMODPS können diese Ansätze einfach und anwenderorientiert implementiert werden.

3.4 Herausforderungen für EMODPS

Grundsätzlich scheint, dass DPS die Flüche der Dynamischen Programmierung umgehen kann, jedoch selbst mit Problemen konfrontiert ist. Der größte Unterschied zwischen den beiden Ansätzen ist, dass DP immer eine optimale Lösung findet, da die Suche nach der besten Lösung über den gesamten Zustand-Aktions-Raum (Kombinationen aus Systemvariablen und Entscheidungen) durchgeführt wird. Es werden quasi alle Varianten an Entscheidungen getestet und die beste Entscheidung gefunden. Im Vergleich dazu sucht DPS nur einen Teil des Zustands-Aktions-Raumes ab und es besteht deshalb keine theoretische Garantie, die optimale Lösung zu finden. Es wird jedoch eine Lösung gefunden, die dem Optimum sehr nahe ist. Ob eine, dem Optimum sehr nahe, Lösung gefunden wird hängt in erster Linie von der Wahl der Funktion ab, die die Strategie abbilden soll. Eine möglichst hohe Flexibilität der Funktion ist gefordert. Besonders wichtig sind außerdem der Informationsgehalt der Input-Daten (I_t) und die Wahl des externen Optimierungsalgorithmus. Dieser muss in der Lage sein, mit nicht linearen Modellen, komplexen Entscheidungsräumen sowie nicht linearen und hoch dimensional Zielfunktionen umgehen zu können.

Eine weitere Herausforderung ist, dass die ermittelte Strategie letztlich nur für den Beobachtungszeitraum und die dort auftretenden Zustands-Aktions-Paare absolut gültig ist. Für Studien, die darauf abzielen die Performanz eines wasserwirtschaftlichen Systems für einen abgeschlossenen Zeitraum zu analysieren (Machbarkeitsstudien, Szenarien-Analysen, etc.), stellt dies kein Problem dar. Mittels EMODPS kann hier eine optimale Lösung approximiert werden.

Insbesondere für den Online-Betrieb von wasserwirtschaftlichen Systemen wirft dies Fragen auf. Wie reagiert die Strategie zum Beispiel auf Zustände, die in der Optimierung nicht aufgetreten sind? Um die erlernte Strategie gegenüber unbeobachteter Zustände robust zu gestalten, könnten die Beobachtungen durch die Erzeugung einer stochastischen Zeitreihe erweitert und für die Optimierung verwendet werden. Die dadurch erweiterten Zustands-Aktions-Paare können dieses Problem jedoch nur lindern. Es werden zwar mehr Zustände beobachtet – theoretisch müsste die Zeitreihe jedoch unendlich lang sein um alle möglichen Zustände in der Optimierung zu berücksichtigen. In der Praxis ist dies natürlich nicht möglich. Es wird davon ausgegangen, dass eine angemessen lange Zeitreihe (z.B. mindestens 50 Jahre) ausreichend viele Zustände abbildet, um eine Strategie für den „Normalfall“ abbilden zu können. Damit die Strategie gegenüber Extremfällen und nicht beobachteter Zustände robust wird muss die Strategie (insbesondere kNN) kritisch analysiert und gegebenenfalls im Nachgang der Optimierung angepasst werden. Grundsätzlich bedarf diese Fragestellung noch weiterführender Untersuchungen.

4 Fallstudie

In der vorgestellten Fallstudie soll der Einfluss hydro-klimatischer Variabilität auf die Performanz des Wasserversorgungssystems der Metropolitan Region Manila untersucht werden. Es werden Betriebsstrategien für einen robusten Reservoirbetrieb vorgeschlagen. Der Fokus der Untersuchung liegt dabei auf dem Angat-Umiray Mehrzwecksystem, das 98% des Wasserbedarfs von Metro Manila deckt und auch der Bewässerung dient. Der optimale Betrieb eines Reservoirs zur Wasserversorgung kann nicht mittels einer Zielfunktion bewertet werden. Grundsätzlich ist wichtig, dass der Bedarf an möglichst vielen Tagen im Jahr gedeckt ist. Kommt es zu einem Versorgungsengpass, darf dieser nicht zu lange andauern und nicht zu stark ausfallen. Hinzu kommt, dass mehrere Nutzer einer Anlage ähnlich viele Ansprüche stellen. Somit handelt es sich beim optimalen Betrieb um ein mehrdimensionales Optimierungsproblem. Geeignete evolutionsbasierte Algorithmen können bis zu zehn Ziele gleichzeitig handhaben. Ein weiterer Vorteil von EMODPS ist, dass es die unkomplizierte Einbindung exogener Information in den Betrieb eines Reservoirs ermöglicht. Ziel ist es zu zeigen, dass das mehrdimensionale Optimierungsproblem mittels EMODPS gelöst werden kann. Weiters soll das „hidden state“ Problem unter Zuhilfenahme von Information, die den hydro-klimatischen Systemzustand (El-Niño-Southern-Oscillation-Index (ENSO)) beschreibt, gelockert werden. Vorläufige Ergebnisse zeigen, dass die optimierte Betriebsstrategie unter Zuhilfenahme von ENSO die Systemzuverlässigkeit verbessert und die Intensität von Versorgungsengpässen verringert – und damit das Risiko von Wasserknappheit reduziert.

5 Literatur

- Bellman, R. (1957). *Dynamic programming*. Princeton, NJ: Princeton University Press.
- Bertsekas, D., & Tsitsiklis, J. (1996). *Neuro-dynamic programming*. Boston, MA: Athena Scientific
- Beyer, H. G., & Sendhoff, B. (2007). Robust optimization - A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, 196(33–34), 3190–3218.
- Castelletti, A., de Rigo, D., Rizzoli, A. E., Soncini-Sessa, R., & Weber, E. (2007). Neuro-dynamic programming for designing water reservoir network management policies. *Control Engineering Practice*, 15(8)
- Doyle, J. C., Francis, B. A., & Tannenbaum, A. (1990). *Feedback Control Theory*
- Deutsches TalsperrenKomitee e.V. Hrsg (2013). *Talsperren in Deutschland*. Springer
- Giuliani, M., Mason, E., Castelletti, A., & Pianosi, F. (2014). Universal approximators for multi-objective direct policy search in water reservoir management problems : a comparative analysis. *Proceedings of the 19th World Congress IFAC*
- Giuliani, M., Ph, D., Castelletti, A., Pianosi, F., Mason, E., Reed, P. M., & Asce, A. M. (2016). Curses , Tradeoffs , and Scalable Management : Advancing Evolutionary Multiobjective Direct Policy Search to Improve Water Reservoir Operations, 142(2).
- Hall, W. A., & Buras, N. (1961). The dynamic programming approach to water resources development. *Journal of Geophysical Research*, 66(2), 510–520.

- Hasinoff, S. W. (2003). Reinforcement Learning for Problems with Hidden State. University of Toronto, Technical Report, 1–18.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement Learning : A Survey, 237–285.
- Koutsoyiannis, D., & Economou, A. (2003). Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems. *Water Resources Research*, 39(6).
- Lehner, B., C. Reidy Liermann, C. Revenga, C. Vörösmart, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J.C. Robertson, R. Rodel, N. Sindorf, and D. Wisser. 2011. High-Resolution Mapping of the World's Reservoirs and Dams for Sustainable River-Flow Management. *Frontiers in Ecology and the Environment*
- Moriarty, D. E., Schultz, A. C., & Grefenstette, J. J. (1999). Evolutionary Algorithms for Reinforcement Learning, 11, 241–276.
- Oliveira, R., & Loucks, D. P. (1997). Operating rules for multireservoir systems. *Water Resources Research*, 33(4), 839.
- Powell, W. (2007). *Approximate dynamic programming: Solving the curses of dimensionality*, Wiley, Hoboken, NJ.
- Puterman, M (2008). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*
- Rippl, W. (1883). The Capacity of Storage-Reservoirs for Water-Supply. *Minutes Proc. Inst. Civ. Eng*, 71, 270–278.
- Rosenstein, M. T., & Barto, A. G. (2001). Robot weightlifting by direct policy search. *IJCAI International Joint Conference on Artificial Intelligence*, 839–844.
- Zatarain Salazar, J., Reed, P. M., Herman, J. D., Giuliani, M., & Castelletti, A. (2016). A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. *Advances in Water Resources*, 92, 172–185.