

RELY – Reliable Reinforcement Learning for Sustainable Energy Systems

Carlotta Tubeuf, Anton Maly, Hugo Götsch und René Hofmann

#vgbe energy - All rights reserved / Alle Rechte vorbehalten - © #

Im Zuge der Energiewende und des steigenden Anteils fluktuierender erneuerbarer Energiequellen rückt die Flexibilität von Speichertechnologien zunehmend in den Fokus – insbesondere die Rolle von Pumpspeicherkraftwerken (PSKW) als verlässliche und bewährte Speicherlösung. Damit diese künftig nicht nur zur Energiespeicherung, sondern auch zur Bereitstellung schneller Systemdienstleistungen beitragen können, sind innovative und adaptive Steuerungskonzepte erforderlich. Genau hier setzt das Forschungsprojekt RELY – Reliable Reinforcement Learning for Sustainable Energy Systems an, das von der Österreichischen Forschungsförderungsgesellschaft (FFG) im Rahmen des Programms „AI for Green“ gefördert wird. Ziel des Projekts ist es, Methoden

des Reinforcement Learning (RL) für eine automatisierte, präzise und dynamisch anpassbare Prozessführung in PSKW nutzbar zu machen. Als erster Use Case wird der Pumpenstart einer Pumpturbine betrachtet – ein komplexer Vorgang mit zahlreichen gekoppelten Komponenten. Erste Simulationsergebnisse zeigen vielversprechendes Potenzial für eine beschleunigte und effizientere Steuerung durch RL.

1 Einleitung

Um das Ziel der Klimaneutralität bis 2050 zu erreichen, wie es in der Europäischen Union angestrebt wird, ist ein umfassender Ausbau erneuerbarer Energiesysteme unerlässlich. Insbesondere Wind- und Photovoltaikanlagen werden ihre Stromproduktion in den kommenden Jahrzehnten voraussichtlich um ein Vielfaches steigern. Die mit diesen Energieerzeugungsanlagen verbundene hohe Volatilität stellt jedoch neue Herausforderungen für die Netzstabilität dar: Lastspitzen und Einspeiseschwankungen müssen zunehmend flexibel ausgeglichen werden. Dies erfordert eine tiefgreifende Transformation der europäischen Energieinfrastruktur hin zu einem hochgradig anpassungsfähigen und intelligenten Energiesystem [Komarnicki et al., 2021]. Eine Schlüsselrolle übernimmt dabei die Flexibilisierung von Energiespeichern – insbesondere Pumpspeicherkraftwerken (PSKW), da sie derzeit noch als einzige großtechnisch verfügbare, wirtschaftlich tragfähige Speicherlösung gelten [Blakers et al., 2021].

Um diese Flexibilität zu erreichen, können Machine Learning Methoden zukünftig zentrale Aufgaben bei der Steuerung, Koordination und Stabilisierung im Energiesystem übernehmen [Ibrahim et al., 2020]. Im Fokus des Projekts RELY – Reliable Reinforcement Learning for Sustainable Energy Systems steht eine spezielle Form von Machine Learning: das Reinforcement Learning (RL). RL-Agenten lernen durch Interaktion mit

ihrer Umgebung optimale Strategien zur Prozesssteuerung. RL ist damit in der Lage, auch bei unvollständiger Datenlage auf veränderte Umgebungsbedingungen flexibel zu reagieren. Durch diese präzise Prozesssteuerung kann RL bislang ungenutzte Effizienzpotenziale erschließen und die Leistungsfähigkeit von Energiesystemen deutlich steigern. Das macht RL besonders geeignet für komplexe, dynamische Systeme wie PSKW.

Ein zentrales Hindernis für die praktische Anwendung ist bislang die mangelnde Nachvollziehbarkeit und Vertrauenswürdigkeit von Machine Learning Methoden, um sicherheitskritische Prozesse zu steuern. Um diesem Risiko zu begegnen, untersuchen wir in RELY einen mehrstufigen Ansatz: Zunächst wird der Algorithmus an einem virtuellen Abbild des Systems vortrainiert und erst anschließend auf das reale System übertragen, unterstützt durch eine Digital-Twin-Plattform. Diese Vorgehensweise ermöglicht es dem Lernalgorithmus, seine Strategie an die tatsächlichen Bedingungen anzupassen und den Zielprozess optimal zu steuern, ohne dabei sicherheitskritische Bereiche des Aktionsraumes im realen Betrieb zu durchlaufen. Darüber hinaus beschleunigt das Vortrainieren in einer Simulationsumgebung den Trainingsprozess erheblich und spart somit wertvolle Zeit und Ressourcen [Tubeuf et al., 2023].

Als Anwendungsfall zur Demonstration des Potentials von RL zu Flexibilisierung von Energiesystemen dient eine labormaßstäbliche Pumpturbine. Diese weist zentrale Eigenschaften fortschrittlicher Energieerzeugungsanlagen auf – wie eine hohe Komplexität, Sicherheitsrelevanz und schnelle Reaktionszeiten – und ist somit ideal geeignet, um den Einsatz von RL in realitätsnahen Szenarien zu erproben. Ziel ist es, kritische Prozesse wie die Start- und Stoppvorgänge, sowie die Umschaltvorgänge zwischen Pump- und Turbinenbetrieb durch RL effizienter und sicherer zu gestalten.

Autoren

Carlotta Tubeuf
Anton Maly
René Hofmann
Institut für Energietechnik und
Thermodynamik
TU Wien
Wien, Österreich

Hugo Götsch
TIWAG-Tiroler Wasserkraft AG
Innsbruck, Österreich



Die Mitglieder des Konsortiums wurden gezielt ausgewählt, um alle erforderlichen Kompetenzbereiche abzudecken, die notwendig sind, um die Projektziele zu erreichen, das Innovationspotenzial voll auszuschöpfen und einen zügigen Transfer der Ergebnisse in industrielle Anwendungen sicherzustellen. Dies ermöglicht es uns, unsere Methode über alle Technologiereifegrade hinweg zu entwickeln – von theoretischen Konzepten über die digitale Integration, die praktische Umsetzung und Automatisierung im realen Anwendungsfall bis hin zur Übertragbarkeit und Anwendbarkeit auf Steuerungssysteme von Anlagenbetreibern, siehe Bild 1.

- TU Wien, Institut für Energietechnik und Thermodynamik (IET): Koordination des Projekts, Entwicklung der digitalen Zwillinge und Implementierung der RL-Methoden. Betrieb der Versuchsanlage der reversiblen Pumppturbine zur Validierung der Algorithmen und Entwicklung des Messkonzepts.
- TU Wien, Institut für Information Systems Engineering: Entwicklung und Absicherung der theoretischen Grundlagen der RL-Algorithmen, insbesondere durch Off-Policy Evaluierung.
- evon GmbH: Umsetzung der Automatisierungslösungen, Schnittstellenmanagement.
- TIWAG – Tiroler Wasserkraft AG: Bereitstellung von Anwendungswissen und Evaluierung realer Betriebsbedingungen.
- vgb energy e.V.: Verankerung des Projekts in der Industrie, Dissemination und Stakeholder-Dialog auf internationaler Ebene.

3 Methodik: Reinforcement Learning für die Prozesssteuerung

forcement Learning nicht aus bestehenden Datensätzen, sondern durch direkte Interaktion mit der Umgebung. Dabei entwickelt der Lernalgorithmus eine optimale Entscheidungsstrategie, die sogenannte Policy, die das Verhalten des Agenten in verschiedenen Situationen bestimmt [Sutton and Barto, 2018].

The diagram illustrates the Reinforcement Learning loop. It features two main components: the **Agent** and the **Environment**, each represented by a blue rounded rectangle. To the left of the Agent is a small robot icon, and to the right of the Environment is a robotic arm icon. Three curved arrows connect the Agent and Environment in a cycle: a top arrow labeled **Action** points from the Agent to the Environment; a bottom arrow labeled **Reward** points from the Environment to the Agent; and a left arrow labeled **State** points from the Environment back to the Agent.

Zu jedem Zeitpunkt t wählt der Agent eine Aktion (Action) A_t , die zu einer Änderung in der Umgebung (Environment) führt. Daraufhin wird dem Agenten der neue Zustand (State) S_t sowie eine Belohnung (Reward) R_t zurückgemeldet. Der Reward gilt als Maß dafür, wie gut dieser Zustand in Hinblick auf die Erreichung des Lernziels ist. Auf Basis dieser Rückmeldung entscheidet der Agent über die nächste Aktion. Durch Wiederholung dieser Abfolge lernt der Agent, States optimal mit Actions zu verknüpfen, um über die Zeit hinweg die kumulierte Belohnung

Für die Steuerung von Prozessen in sicherheitskritischen Infrastrukturen wie PSKWs mittels RL, ergeben sich jedoch zwei zentrale Herausforderungen für das Trainieren des RL Agenten: Erstens kann das Training tausende von Durchläufen benötigen, bis der Agent durch Ausprobieren aller möglichen Aktionen und Abfolgen zu einer optimalen Policy findet. Und zweitens trifft der Agent während des Trainings nicht nur logische und sinnvolle Actions, was zu erheblichen Sicherheitsrisiken bei einer Anwendung in realen Anlagen führen kann.

4 Use Case: Startvorgang im Pumpbetrieb

Um in Zukunft flexibler auf Netzanforderungen reagieren zu können, wird der Wechsel zwischen den Betriebszuständen – Pumpbetrieb, Turbinenbetrieb, Phasenschieberbetrieb und Stillstand – in Pumpspeicherkraftwerken (PSKW) zunehmend wichtiger [Nikolaos et al., 2023]. Ziel ist es nicht nur, überschüssige Energie aus dem Stromnetz



zu speichern und in Zeiten erhöhten Verbrauchs wieder bereitzustellen, sondern auch, frequenzstabilisierende Systemdienstleistungen effizient zur Verfügung zu stellen. Damit diese Übergänge möglichst schnell, sicher und effizient erfolgen können, ohne beispielsweise unzulässige Druckschwankungen zu verursachen, ist eine präzise und dynamische Steuerung der beteiligten Prozesse erforderlich [Hirth, 2016]. RL bietet hierfür vielversprechende Potenziale, da es durch seine Echtzeitfähigkeit komplexe Steuerungsaufgaben adaptiv und autonom übernehmen kann [Tubeuf et al., 2024].

Im Projekt RELY wird der Pumpenstart als erster Use Case für einen optimierten Betriebswechsel untersucht. Dieser Vorgang eignet sich besonders gut zur Demonstration der Möglichkeiten einer RL-basierten Prozesssteuerung, da er mehrere kritische und dynamisch gekoppelte Systemkomponenten umfasst, die exakt aufeinander abgestimmt werden müssen, um einen stabilen und schnellen Übergang in den Pumpbetrieb zu gewährleisten.

In Bild 4 sind alle für den Pumpenstart relevanten Komponenten dargestellt. Um bewerten zu können, wie sich die Steuerung des Pumpenstarts mittels RL im Vergleich zur herkömmlichen Steuerung verhält, wurde eine manuelle Abfolge der Prozessschritte als Ausgangs- und Vergleichsbasis definiert. Als Vorbild dient der Ausblasvorgang im Kraftwerk Kühtai 1 der TIWAG. Das zugehörige Ablaufdiagramm ist in Bild 5 abgebildet.

Ausgehend vom Stillstand des Maschinensatzes werden beim Pumpenstart zunächst

die Leitschaukeln geschlossen. Anschließend wird Luft in den Konus eingeblasen, um das Wasser aus dem Laufrad zu verdrängen. Dadurch kann das Laufrad mit einem Anfahrmotor auf die Synchronisationsdrehzahl beschleunigt werden, wobei das aufgenommene Drehmoment – und damit die benötigte Leistung – durch die Rotation in Luft minimiert wird. Sobald die Pumpturbine synchronisiert und ans Netz angeschlossen ist, werden der Kugelhahn und die Leitschaukeln geöffnet. Das Wasser wird nun vom Unterwasser- zum Oberwasserreservoir gepumpt.

Für die Steuerung des Vorgangs mittels RL wurden folgende States und Actions festgelegt:

States (Zustandsgrößen)

- Ventilstellung Kugelhahn
- Ventilstellung Luftventil
- Leitschaukelstellung
- Drehzahl
- Drehmoment
- Durchfluss
- Geodätische Höhe
- Luftvolumenstrom
- Wasserlevel in Konus
- Zeit

Diese States repräsentieren den aktuellen Systemzustand zu jedem Zeitschritt und liefern dem RL-Agenten die erforderlichen Informationen, um fundierte Steuerentscheidungen zu treffen.

Actions (Steuerbefehle)

- Kugelhahn öffnen/schließen
- Luftventil öffnen/schließen

- Leitschaukelöffnung
- Drehzahl mit konstanter Beschleunigung hochfahren

Diese Aktionen ermöglichen dem RL-Agenten, sämtliche relevanten Aktoren zur Durchführung des Pumpenstarts flexibel und optimal zu steuern.

Um die optimale Policy für die Steuerung des Pumpenstarts mittels Reinforcement Learning (RL) zu erlernen, ist eine präzise definierte Belohnungsfunktion (Rewardfunktion) erforderlich. Diese berücksichtigt sowohl ökonomische als auch technische Zielgrößen. Die Rewardfunktion ist als gewichtete Summe mehrerer Terme formuliert, die positive Anreize (Belohnungen) und negative Anreize (Strafen) abbilden. Sie dient der Gesamtbewertung des Steuerungsverhaltens und umfasst unter anderem:

- Bestrafung für Energieverlust, solange die Pumpturbine nicht am Netz ist
- Belohnung bei erreichtem Zieldurchfluss
- Kosten für die Synchronisation durch den Anfahrmotor
- Anteilige Abnutzung mechanischer Komponenten
- Stromkosten für die Luftbereitstellung
- Verschleißkosten durch Ventilbetätigungen

Alle Terme werden durch geeignete Gewichtung auf die gemeinsame Einheit Euro normiert und anschließend summiert. Ziel des RL-Agenten ist es somit, die Gesamtkosten des Pumpenstarts zu minimieren und gleichzeitig die Betriebseffizienz zu maximieren.

Im Rahmen der Simulation wurden mehrere RL-Agenten mit unterschiedlichen Lernalgorithmen darauf trainiert, eine optimale Steuerungsstrategie für den Pumpenstart zu entwickeln. Als besonders vielversprechend erwies sich dabei der PPO-Agent (Proximal Policy Optimization), der im Vergleich zu den übrigen getesteten Ansätzen die stabilsten und effizientesten Lernfortschritte zeigte. Erste Ergebnisse aus dem Simulationsmodell deuten darauf hin, dass dieser Agent in der Lage ist, eine eigenständige Steuerstrategie zu erlernen, die den Übergang in den Pumpbetrieb nicht nur zuverlässig, sondern auch schneller als das manuelle Abfollgeschema vollzieht. Der direkte Vergleich legt nahe, dass der RL-Agent durch eine präzise und dynamisch angepasste Aktorsteuerung den Startprozess beschleunigt und somit das Potenzial besitzt, konventionelle Steuerkonzepte hinsichtlich Effizienz und Reaktionsgeschwindigkeit zu übertreffen.

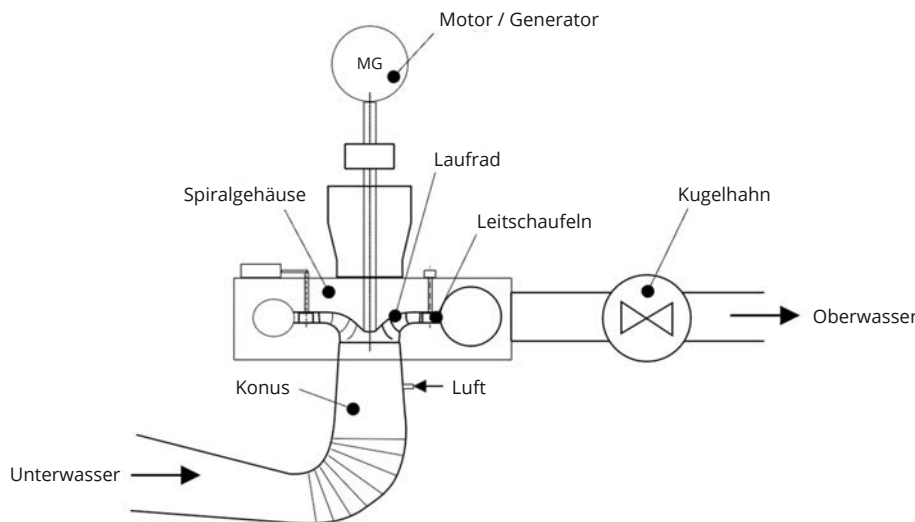


Bild 4. Querschnitt einer reversiblen Pumpe mit für den Pumpenstart wichtigen Komponenten.

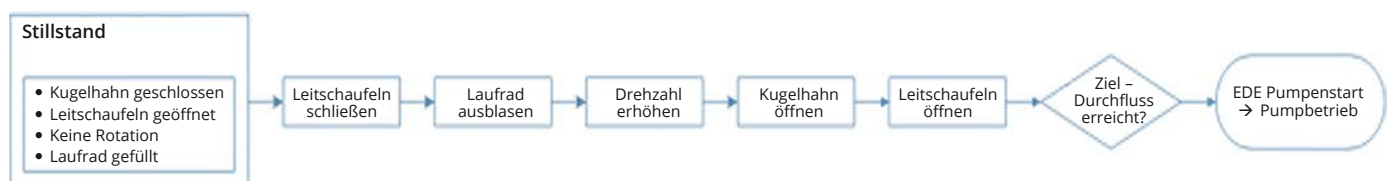


Bild 5. Generelles Fließschema eines Pumpenstarts.

5 Ausblick

Ob sich die vielversprechenden Ergebnisse des Trainings im Simulationsmodell auch im realen Einsatz an der Modell-Pumpturbine im Labor des Instituts für Energiesysteme und Thermodynamik (IET) der TU Wien bestätigen, wird sich in den kommenden Testreihen zeigen.

Dafür erfolgt die Implementierung des Automatisierungssystems der Firma evon im Labor für Strömungsmaschinen des IET an der TU Wien, um reale Versuchsreihen durchzuführen. Dabei sollen Sicherheitsbereiche definiert werden, in denen die manuelle Steuerung bei Bedarf die Kontrolle vom RL-Agenten übernimmt. Anschließend ist die Übertragung des RL-Agenten auf den Betrieb der Modell-Pumpturbine vorgesehen. Ziel ist es, den potenziellen Mehrwert des RL-Ansatzes für eine erhöhte Betriebsflexibilität systematisch zu evaluieren und in weiterer Folge die Steuerung weiterer Prozesse, wie den Turbinenstart oder den kompletten Betrieb, in Pumpspeicherkraftwerken RL-basiert zu erweitern.

Parallel wird das Konzept der Off-Policy Evaluation zur quantitativen Risikobewertung eingesetzt. Damit kann abgeschätzt werden, wie wahrscheinlich es ist, dass ein in der Simulation erlerntes Verhalten auch im realen System sicher und optimal bleibt – ohne dass das System tatsächlich eingesetzt werden muss.

Langfristig soll diese Methodik als Grundlage für die Steuerung komplexer, volatiler Energiesysteme dienen – von Wasserkraft über Batteriespeicher bis hin zu virtuellen

Kraftwerken. Damit leistet das Projekt einen essenziellen Beitrag zur Realisierung eines sicheren, effizienten und klimaneutralen Energiesystems der Zukunft.

Bibliographie

- Blakers, A., Stocks, M., Lu, B., Cheng, C., 2021. A review of pumped hydro energy storage. *Prog. Energy* 3, 022003. <https://doi.org/10.1088/2516-1083/abeb5b>.
- Hirth, L., 2016. *The benefits of flexibility: The value of wind energy with hydropower*. *Applied Energy* 181, 210–223. <https://doi.org/10.1016/j.apenergy.2016.07.039>.
- Komarnicki, P., Kranhold, M., Styczynski, Z.A., 2021. *Flexibilität eines Gesamtenergiesystems (GES)*, in: Komarnicki, P., Kranhold, M., Styczynski, Z.A. (Eds.), *Sektorenkopplung – Energetisch-nachhaltige Wirtschaft der Zukunft: Grundlagen, Modell und Planungsbeispiel eines Gesamtenergiesystems (GES)*. Springer Fachmedien, Wiesbaden, pp. 133–163. https://doi.org/10.1007/978-3-658-33559-5_5.
- Nikolaos, P.C., Marios, F., Dimitris, K., 2023. A Review of Pumped Hydro Storage Systems. *Energies* 16, 4516. <https://doi.org/10.3390/en16114516>.
- Sutton, R.S., Barto, A., 2018. *Reinforcement learning: An introduction*, Second edition. ed, *Adaptive computation and machine learning*. The MIT Press, Cambridge, Massachusetts; London, England.
- Tubeuf, C., aus der Schmitt, J., Hofmann, R., Heitzinger, C., Birkelbach, F., 2024. *Improving Control of Energy Systems With Reinforcement Learning: Application to a Reversible Pump Turbine*. Presented at the ASME 2024 18th International Conference on Energy Sustainability collocated with the ASME 2024 Heat Transfer Summer Conference and the ASME 2024 Fluids Engineering Division Summer Meeting, American

Society of Mechanical Engineers Digital Collection. <https://doi.org/10.1115/ES2024-122475>.

Tubeuf, C., Birkelbach, F., Maly, A., Hofmann, R., 2023. *Increasing the Flexibility of Hydropower with Reinforcement Learning on a Digital Twin Platform*. *Energies* 16, 1796. <https://doi.org/10.3390/en16041796>.

Abstract

RELY – Reliable Reinforcement Learning for Sustainable Energy Systems

In the wake of the energy transition and the increasing share of fluctuating renewable energy sources, the flexibility of storage technologies is becoming increasingly important – in particular the role of pumped storage power plants (PSPPs) as a reliable and proven storage solution. Innovative and adaptive control concepts are needed to ensure that these plants can contribute not only to energy storage but also to the provision of fast system services in the future. This is precisely where the RELY – Reliable Reinforcement Learning for Sustainable Energy Systems research project, funded by the Austrian Research Promotion Agency (FFG) as part of the 'AI for Green' programme, comes in. The aim of the project is to make reinforcement learning (RL) methods usable for automated, precise and dynamically adaptable process control in PSKWs. The first use case to be considered is the pump start-up of a pump turbine – a complex process involving numerous interconnected components. Initial simulation results show promising potential for accelerated and more efficient control through RL.



vgbe-Standard VGBE-S-197-00-2024-04-DE

Statische und quasistatische Dichtungen

vgbe-Standard VGBE-S-197-00-2024-04-DE. 58 S., 23 Abb., DIN A4 (2024), ISBN 978-3-96284-320-5 (Print, Deutsch), ISBN 978-3-96284-321-2 (E-Book, Deutsch)
Dieser vgbe-Standard liegt nur in Deutscher Sprache vor.
This vgbe-Standard is only available in German.
Preis für vgbe-Mitglieder* 135,- €, Nichtmitglieder 195,- €, + Versand und USt.

In diesem vgbe-Standard werden Hinweise für den Einsatz von Dichtungen und deren Ausführung in Energieanlagen gegeben. Auslegung, Werkstoffauswahl, Prüfung und Handhabung von statischen und quasistatischen Dichtungen werden behandelt.

