Feedback Control Approaches For Restoration of Power Grids From Blackouts

Joseph M. Miller, Hugo N. Villegas Pico, Ian Dobson
Department of Electrical and Computer Engineering
Iowa State University, Ames, IA USA
{jmm1, hvillega, dobson}@iastate.edu

Andrey Bernstein and Bai Cui
Power Systems Engineering Center
National Renewable Energy Laboratory, Golden, CO USA
{Andrey.Bernstein, Bai.Cui}@nrel.gov

***Abstract*—The automated restoration of power systems with variable energy resources is a timely problem to tackle. Automated restoration advice can support operators in deciding on strategic actions to restore power grids from a blackout with a mix of conventional and renewable generation resources. To this end, this paper frames the restoration process of power grids with solar resources as a nonlinear dynamic model with algebraic constraints in discrete time which is steered by feedback control loops. We discuss two feedback-control strategies based on greedy and reinforcement learning algorithms, and contrast their performance with restoration plans generated by a mixed-integer linear program. We found that the reinforcement learning algorithm infers restoration actions faster than the greedy one. However, the tuning process of the reinforcement learning parameters is slower than for the greedy one.**

***Index Terms*—Restoration, machine learning, optimization.**

I. Introduction

A classical resilience problem in power engineering is the restoration of an electric power grid from a blackout [1], [2]. A restoration process uses black-start generating units (hydro or gas powered [3], [4]) to: (i) energize transmission paths, (ii) supply demand, and (iii) start up non-black-start power plants. When a blackout is declared, a step-by-step restoration plan is deployed to quickly re-energize the grid [5, p. 183]. During its execution, power system operators have the difficult task of analyzing whether the restoration process is progressing well [2], [5], [6]. They *perceive* this by monitoring grid variables, e.g., voltage magnitudes, frequency, as well as generation and transmission loading [6].

The classical focus of the restoration problem is on deciding cranking paths to energize critical loads and next-start units [1], [7]. This is because the primary energy resources of conventional generation are assumed firm, e.g., gas and water supply [1], [4]. At present, the decision-making for restoration is challenged by the variability of renewable energy resources such as wind and solar [8]–[10]. In particular, operators are now confronted to *adapt* restoration strategies on the fly to compensate for the variability of non-conventional generation. Hence, advancing automated restoration methods is timely.

Automated restoration techniques have been researched since the advent of digital computers [11]. Present approaches for restoration resort to optimization algorithms [7], [10], [12], decision trees [6], and reinforcement learning [13]–[19], to cite a few. In these methods, decisions are incentivized by either tracking the number of energized transmission lines, the amount of energized load, or by testing whether grid variables are within specifications [6], [17]. Other relevant metrics from resilience focus on economics [18], [20], [21].

A contemporary challenge to the design of feedback controllers for restoration pertains to the complexity of representing the re-energization process of power grids. At present, simulating *one* full restoration process using electromagnetic transient programs or positive-sequence approaches (which are computationally lighter) can take several hours [22]. Notably, the variability of the primary energy resources like solar and wind power are omitted because they are assumed invariant during transients [23]–[25]. Another problem for designing restoration controllers pertains to determining a set of performance metrics to stimulate automatic restoration decisions. The purpose, for example, is to steer the system states towards desired targets [26, Ch. 5].

To address the aforementioned challenges, we report the following contributions: (i) An abstract model of the power system restoration process in a minute time scale represented by a set of nonlinear coupled discrete algebraic equations which incorporate the variability of solar resources. This model is computationally light, hence useful to iteratively design restoration feedback controls. A limitation is the omission of electromechanical and electromagnetic transients. (ii) A set of restoration performance metrics to evaluate the resilience of grids in the context of blackout recovery. These metrics model in a low dimensional space the status of voltage magnitudes, breaker statuses, amount of load served, equity in load energization, and energy storage level, to name a few. These serve to stimulate a controller to make decisions so that energy transmitted from generation to the grid loads is maximized. (iii) A computer implementation of greedy and reinforcement-learning control strategies to steer the restoration process by employing the aforementioned restoration performance metrics. We showcase these contributions in a power system example with photovoltaic and concentrating solar assets. This paper has objectives to meet restoration requirements [27] and mitigate likely impacts from blackouts.

The paper is organized as follows. In Section II, we develop details of the restoration process and the modeling approach. In Section III, we propose a set of blackout restoration (or resilience) performance metrics. In Section IV, we design

state feedback controls employing greedy and reinforcement learning approaches. In Section V, we present an illustrative case study that showcases the contributions of the paper. The paper concludes in Section VI.

II. Restoration Process Model

The restoration process is modeled with nonlinear dynamics and algebraic constraints in discrete time:

$$\boldsymbol{x}_{k+1} = \mathscr{F}(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{a}_k, \boldsymbol{d}_k, \boldsymbol{\xi}_k) \tag{1}$$

$$\boldsymbol{0} = \mathscr{G}(\boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1}) \tag{2}$$

$$\boldsymbol{r}_k = \mathscr{R}(\boldsymbol{x}_k, \boldsymbol{y}_k)\,. \tag{3}$$

The discrete times are t_k with discrete time index $k = 0, 1, 2, \ldots, K-1$. The time step $t_{k+1} - t_k$ is assumed constant for all k. The vector $\boldsymbol{x}_k = \boldsymbol{x}(t_k) \in \mathbb{R}^{n_x}$ contains both continuous and discrete states. The vector $\boldsymbol{y}_k = \boldsymbol{y}(t_k) \in \mathbb{R}^{n_y}$ contains continuous output variables. The restoration is steered by the control input $\boldsymbol{a}_k = \boldsymbol{a}(t_k) \in \mathbb{R}^{n_a}$. The exogenous disturbances are modeled with the vector $\boldsymbol{d}_k = \boldsymbol{d}(t_k) \in \mathcal{D} \subset \mathbb{R}^{n_d}$. The vector $\boldsymbol{\xi}_k = \boldsymbol{\xi}(t_k) \in \mathcal{E} \subset \mathbb{R}^{n_\xi}$ models impacts of uncertainties on the system dynamics. Equation (3) defines a vector of state-dependent restoration performance metrics $\boldsymbol{r}_k = \boldsymbol{r}(t_k) \in \mathbb{R}^{n_r}$, which are defined in Section III. In Section IV-A, these metrics are used to stimulate greedy restoration control actions. They are also used to stimulate reinforcement learning in Section IV-B.

In more detail, the vector $\boldsymbol{x}_k$ describes at time t_k the state of the circuit breakers, the amount of load that has been picked up, the amount of stored energy, and the amount of generation by the solar resources. The vector $\boldsymbol{y}_k$ describes the bus voltage magnitudes and phase angles, real and reactive power flows, and the power injections from the conventional generators. The vector $\boldsymbol{a}_k$ models commands such as to close a circuit breaker, pick up load, or increase the contribution from solar resources. The vector $\boldsymbol{d}_k$ contains time varying inputs such as the power that can be delivered by PV arrays or concentrating solar technologies. The generated power, energized load, and exogenous inputs are perturbed by $\boldsymbol{\xi}_k$.

The restoration process may be controlled by a variety of decision-making approaches that might or might not need state feedback as illustrated in Fig. 1. The action $\boldsymbol{a}_k$ can emerge from a closed-loop controller that uses feedback information of the present state of the system, e.g., greedy techniques and reinforcement-learning agents. Decisions can also be sourced from a list of pre-computed sequential actions (without state feedback), e.g., by solving a mixed-integer linear program (MILP). These controllers may also be stimulated by a forecast $\widehat{\boldsymbol{d}}_k = \widehat{\boldsymbol{d}}(t_k) \in \widehat{\mathcal{D}} \subset \mathbb{R}^{n_d}$. Typically, $\widehat{\boldsymbol{d}}_k \neq \boldsymbol{d}_k$, i.e., the forecast does not match the actual observations.

We consider that power system operators execute one action at a time during a restoration process, e.g., see [5, p. 183]. Here, each action, $\boldsymbol{a}_k$, belongs to a set of admissible actions:

$$\mathcal{U} = \{\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \ldots, \boldsymbol{u}^{(n_u)}\} \subset \mathbb{R}^{n_a}\,. \tag{4}$$

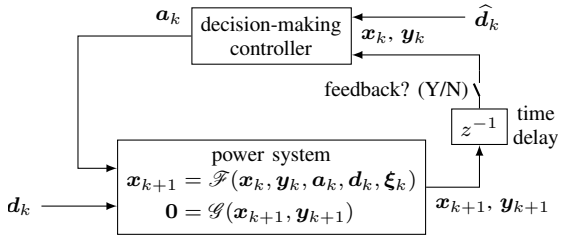


Fig. 1. Restoration controller with and without feedback.

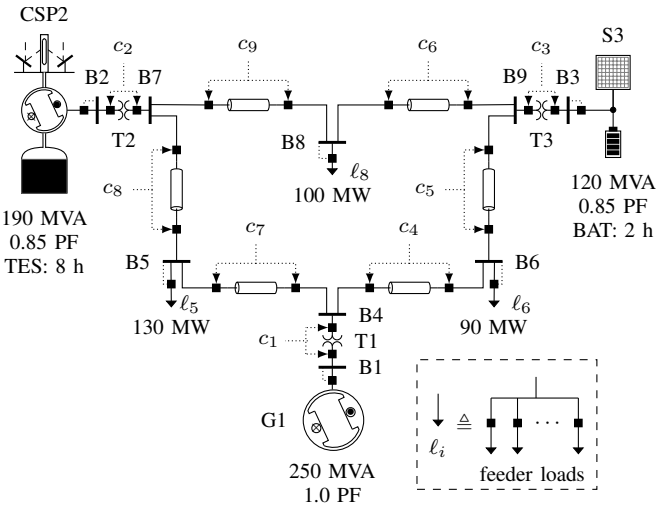


Fig. 2. Modified WSCC 9-bus power system. Closed breakers are filled black. A load ℓ_i, $i \in \{5, 6, 8\}$, models the aggregate impact of various feeder loads.

An admissible action vector with all zero entries means doing nothing. The other admissible action vectors have one non-zero entry which can be ± 1. This means that at most one action is applied to the system at a time; for example, to close a circuit breaker.

We next explain the modeling of the restoration process with the power system in Fig. 2. This power system, which will be the subject of analysis in Section V, has $n_b = 9$ buses (B1–B9), $n_g = 3$ power plants (B1–B3), $n_l = 9$ transmission branches (e.g., {B7,B8}), and $n_\ell = 3$ loads (connected to B5, B6, and B8). All transmission branches have associated circuit breakers for their energization which are commanded by the binary control variables c_1 to c_9. The black-start unit is connected to bus B1 and is assumed to be hydro powered. Further, bus B2 contains a concentrating solar power (CSP) plant with thermal energy storage (TES) which can continuously produce 160 MW (rated power) for eight hours. Bus B3 has a PV power plant with battery energy storage capacity to generate 110 MW for two hours.

At the beginning of restoration, there is no transfer of electricity from generation to the system loads via the transmission system; all circuit breakers are open and only B1 is energized. During restoration, grid operators decide on which branches and generation assets to energize and the amount of load to pick up [5].

A. Power Flow Model

This subsection constructs $\mathscr{G}$ of (2) which is in essence a classical power flow calculation. From Fig. 2, we consider a bulk transmission system with a set of n_b buses:

$$\mathcal{B} = \{\mathrm{B}1, \mathrm{B}2, \ldots, \mathrm{B}n_b\}\,. \tag{5}$$

Per-unit voltage magnitudes and phase angles at all buses are:

$$\boldsymbol{v} = [v_1, v_2, \ldots, v_{n_b}]^\top \text{ and } \boldsymbol{\theta} = [\theta_1 = 0, \theta_2, \ldots, \theta_{n_b}]^\top\,. \tag{6}$$

A power plant (conventional or hybrid) connects to a bus in:

$$\mathcal{B}_g = \{\mathrm{B}1, \mathrm{B}2, \ldots, \mathrm{B}n_g\} \subset \mathcal{B} \tag{7}$$

i.e., the first n_g buses of $\mathcal{B}$ have generators. The first bus has a black-start unit.

All power plants regulate their terminal voltages to follow operator prescribed set-points:[1]

$$\boldsymbol{v}^{\text{gen}} = [v_1^{\text{gen}}, v_2^{\text{gen}}, \ldots, v_{n_g}^{\text{gen}}]^\top\,. \tag{8}$$

The vector of active power supplied at all terminals of the generation resources (other than the black-start unit) is:

$$\boldsymbol{p}^{\text{grid}} = [p_2^{\text{grid}}, p_3^{\text{grid}}, \ldots, p_{n_g}^{\text{grid}}]^\top\,. \tag{9}$$

The variable p_1^{grid} does not appear because it is the slack unit/bus in the power flow problem and corresponds to the black-start generator in a physical grid.[2]

In the same vein, we consider n_ℓ loads that can be connected to any bus of $\mathcal{B}$. The active and reactive power consumed by the loads are:

$$\boldsymbol{\ell} = [\ell_1, \ell_2, \ldots, \ell_{n_\ell}]^\top \text{ and } \boldsymbol{q} = [q_1, q_2, \ldots, q_{n_\ell}]^\top \tag{10}$$

respectively. Here, $q_i = \ell_i\sqrt{(1/\gamma_i)^2 - 1}$, $i = 1, 2, \ldots, n_\ell$ where γ_i is the power factor of load i. The target load to restore for load i is ℓ_i^{target}.

There are n_l transmission branches with a circuit breaker at each end.[3] For simplicity, we assume that both the circuit breakers on a branch have the same status: open or closed. The statuses of all breakers are collected in:

$$\boldsymbol{c} = [c_1, c_2, \ldots, c_{n_l}]^\top\,, \tag{11}$$

where the branch j breakers are either both open ($c_j = 1$) or both closed ($c_j = 0$). Finally, the active powers flowing through the branches are:

$$\boldsymbol{p}^\dagger = [p_1^\dagger, p_2^\dagger, \ldots, p_{n_l}^\dagger]^\top \tag{12}$$

where $p_j^\dagger \in [-p_j^{\dagger\text{rating}}, p_j^{\dagger\text{rating}}]$.

[1] If a generating unit surpasses its rated current limits, regulation of terminal voltages could be challenging. Power converters have hard current control limits that constrain terminal currents at rated levels for continuous operation. On the other hand, synchronous generation may disconnect because of relatively high stator winding temperature.

[2] A black-start resource typically operates in isochronous-governor mode to fix grid frequency at its rated value [3]. Hence, load pick-up is only performed by the black-start resource even if other generators are connected in parallel. Generation levels of other units are adjusted using their set-points.

[3] A transmission branch is either a power transformer or transmission line.

Overall, we construct a power flow problem from (8), (9), (10), and the associated parameters with the objective of calculating p_1^{gen} and a subset of variables in (6), as well as all the bus angles and branch flows. We clarify that this power flow problem is special because all buses are isolated at the beginning of the restoration process, and later they are energized dynamically and their types are changed. Also, the transmission topology depends on the statuses of the circuit breakers that constrain which buses in the system are energized. The magnitude of load that can be changed at each restoration step is commanded dynamically. These particularities are discussed next.

B. Circuit Breaker Dynamics

The dynamics of the circuit breakers on the j-th line are:

$$c_j(t_{k+1}) = c_j(t_k) \oplus c_j^*(t_k) \tag{13}$$

where $c_j(t_k) \in \{0, 1\}$, $\oplus$ denotes *exclusive or*, and $c_j^*(t_k) \in \{0, 1\}$ is a command to flip the breaker state. The exclusive or implements 'close if open' and 'open if closed' with a single breaker control command $c_j^*(t_k)$.

C. Bus Type Selection

Each bus is PQ, PV or a swing type at each timestep t_k. The type of bus r is selected at each timestep t_k according to

PQ bus type	if there is a closed line joining bus r to bus s and $v_s(t_k) > 0$.
PV bus type	if bus r has a generator and $v_r(t_{k-1}) \in [0.9, 1.1]$.
swing type bus	otherwise

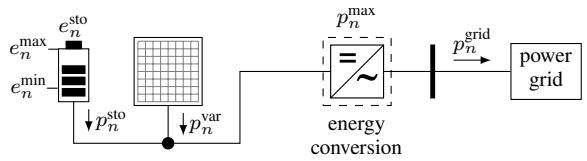
All buses are considered swing type at the beginning of the black-start process, i.e., at $t = t_0$ because all transmission branches are open. The voltage set-point for each isolated bus is zero other than for the bus with a black-start unit.

D. Load Energization Dynamics

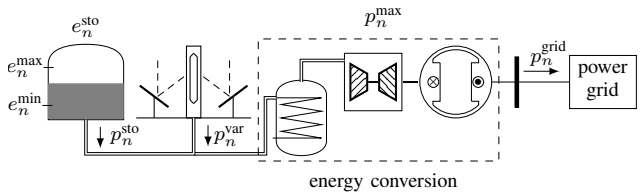
We model the dynamics of load i with:

$$\ell_i(t_{k+1}) = \begin{cases} \ell_i(t_k) + \kappa_i'\ell_i^*(t_k) & \text{if } v_i(t_k) \in [0.9, 1.1] \\ 0 & \text{if } v_i(t_k) \notin [0.9, 1.1]\,. \end{cases} \tag{14}$$

Here, $\ell_i^*(t_k) \in \{-1, 0, 1\}$ is a discrete control command to increment or decrement load and $\kappa_i' = \kappa_i(1 + \xi_{\ell_i})$ with $\xi_{\ell_i} \in [-\overline{\xi}_{\ell_i}, \overline{\xi}_{\ell_i}]$ sampled from a uniform distribution at each time t_k to represent an unknown-but-bounded load pick-up. We note here that transmission operator's command to pick up or reject load is executed by a distribution operator by energizing or de-energizing feeder circuits, which is not modeled here. In (14), we consider that the load can jump to zero if the bus voltage, v_i, is outside prescribed limits by NERC [28, p. 18]. Notably, voltage protection relays can direct the opening of substation circuit breakers to de-energize all connected loads.



(a) hybrid solar PV power plant



(b) hybrid concentrating solar power plant

Fig. 3. Solar power plants with converter and battery storage as well as with steam turbine and thermal energy storage.

E. Power Plant Dynamics

We model here the quasi-steady-state response of a power plant to sequential restoration commands. We first model the dynamics of energy stored by the n-th power plant as any of the ones in Fig. 3. These diagrams illustrate details of non-conventional power plants of Fig. 2 and the variables described in this section. In particular:

$$e_n(t_{k+1}) = e_n(t_k) - \eta_n^{\text{sto}}(p_n^{\text{sto}}(t_k))p_n^{\text{sto}}(t_k)\Delta t_k \tag{15}$$

where $\Delta t_k = t_{k+1} - t_k$ and the efficiency η_n^{sto} of (15) is:

$$\eta_n^{\text{sto}}(p_n^{\text{sto}}) = \begin{cases} \eta_n^{\text{fwd}} & \text{if } p_n^{\text{sto}} \geq 0 \\ 1/\eta_n^{\text{fwd}} & \text{if } p_n^{\text{sto}} < 0 \end{cases}. \tag{16}$$

Here, η_n^{fwd} denotes the efficiency when withdrawing energy from storage. Also, $e_n(t_k) \in [e_n^{\min}, e_n^{\max}]$, $k = 0, 1, \ldots, K$, models a bounded amount of stored energy in the n-th energy storage subsystem at time t_k. The rate at which energy flows into or out of the storage subsystem in (15) (q.v. Fig. 3) per unit time is modeled with:

$$p_n^{\text{sto}}(t_k) = \begin{cases} \frac{1}{\eta_n^{\text{cnv}}} p_n^{\text{grid}}(t_k) - \overline{p}_n^{\text{var}}(t_k) & \text{if } A_n(t_k) \\ \frac{1}{\eta_n^{\text{cnv}}} p_n^{\text{grid}}(t_k) - \overline{p}_n^{\text{var}}(t_k) & \text{if } B_n(t_k) \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

$$A_n(t_k) : \frac{p_n^{\text{grid}}(t_k)}{\eta_n^{\text{cnv}}} > \overline{p}_n^{\text{var}}(t_k) \;\&\; e_n(t_k) \geq e_n^{\min} + \chi_n' \Delta t_k$$

$$B_n(t_k) : \frac{p_n^{\text{grid}}(t_k)}{\eta_n^{\text{cnv}}} < \overline{p}_n^{\text{var}}(t_k) \;\&\; e_n(t_k) < e_n^{\max} - \chi_n' \Delta t_k$$

where A_n and B_n are conditions of the n-th power plant evaluated at t_k. Also, $\eta_n^{\text{cnv}} \in (0, 1]$ is the efficiency of the n-th energy conversion subsystem and:

$$\overline{p}_n^{\text{var}}(t_k) = (1 + \xi_n^{\text{var}})\widehat{p}_n^{\text{var}}(t_k) \tag{18}$$

is the maximum electric power that can be withdrawn from the renewable energy conversion subsystem, e.g., PV array. We note here that this power is modeled as an uncertain deviation from a forecast value $\widehat{p}_n^{\text{var}}(t_k)$ via $\xi_n^{\text{var}} \in [-\overline{\xi}_n^{\text{var}}, \overline{\xi}_n^{\text{var}}]$ which is sampled from a uniform distribution at each t_k. For water and gas resources, $\overline{p}_n^{\text{var}} = 0$.

In (17), $p_n^{\text{grid}}(t_k)$ models the power transferred from a power plant to the grid which satisfies:

$$p_n^{\text{grid}}(t_{k+1}) = \text{Proj}_{\mathcal{P}_n(t_k)} \left\{ p_n^{\text{grid}}(t_k) + \chi_n' p_n^*(t_k) \right\} \tag{19}$$

where $p_n^*(t_k) \in \{-1, 0, 1\}$ is a discrete control command to increase or decrease generation by $\chi_n' = \chi_n(1 + \xi_{p_n})$ where χ_n is the mean and $\xi_{p_n} \in [-\overline{\xi}_{p_n}, \overline{\xi}_{p_n}]$ is sampled from a uniform distribution at each time step k to model uncertain changes in generation.[4] Note that a power plant can supply rated power to the grid only if there is enough stored energy, otherwise, it can only inject the power that can be harvested by the renewable resource. Limits in (19) are enforced via the Euclidean projection operator, $\text{Proj}_{\mathcal{P}}\{\cdot\}$, onto the set:[5]

$$\mathcal{P}_n(t_k) := \begin{cases} [0, p_n^{\max}] & \text{if } e_n(t_k) \geq e_n^{\min} + \chi_n' \Delta t_k \\ [0, \eta_n^{\text{cnv}} \overline{p}_n^{\text{var}}(t_{k+1})] & \text{if } e_n(t_k) \leq e_n^{\min} \end{cases}. \tag{20}$$

Here, $p_n^{\max}$ is the maximum energy conversion rate per unit time. We emphasize that this subsection also applies to power plants fired with fossil fuels or powered by the flow of water. In such cases, $e_n(t_k)$ of (15) is relatively large.

F. Assembly of Restoration Process Model

We assemble the vectors of (1) and (2) as follows:

$$\begin{aligned} \boldsymbol{x}_k &= [\boldsymbol{c}_k; \boldsymbol{\ell}_k; \boldsymbol{e}_k; \boldsymbol{p}_k^{\text{grid}}], \text{ from (13), (14), (19), (15)} \\ \boldsymbol{y}_k &= [p_{1,k}^{\text{grid}}; \boldsymbol{v}_k; \boldsymbol{\theta}_k; \boldsymbol{p}_k^{\dagger}], \text{ from (6), (12)} \\ \boldsymbol{u}_k &= [\boldsymbol{c}_k^*; \boldsymbol{\ell}_k^*; \boldsymbol{p}_k^*], \text{ from (13), (14), (15)} \\ \boldsymbol{d}_k &= \overline{\boldsymbol{p}}_k^{\text{var}}, \text{ from (17), (20)} \\ \boldsymbol{\xi}_k &= [\boldsymbol{\xi}_{\ell k}; \boldsymbol{\xi}_{pk}; \boldsymbol{\xi}_k^{\text{var}}], \text{ from (14), (18), (19)} \end{aligned}$$

III. Restoration Performance Metrics

We present a set of restoration (or resilience) performance metrics. The set of metrics is a low-dimensional representation of the status of the grid. The purpose of these metrics is to design and/or deploy state feedback controllers and to quantify the restoration performance for operator awareness. We construct $\mathscr{R}$ of (3) to output:

$$\boldsymbol{r} = [r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8]^\top \tag{21}$$

which is calculated at each t_k from $\boldsymbol{x}_k$ and $\boldsymbol{y}_k$ of (1) and (2).

A. Restored Branches

We describe the fraction of branches that are closed at a particular point of time with:

$$r_1 := \frac{1}{n_l} \sum_{j=1}^{n_l} c_j \,. \tag{22}$$

Recall that c_j of (11) indicates whether the j-th branch is open or closed. This metric is suited to stimulate the closure of breakers by a restoration controller. If breakers are all open or all closed, $r_1 = 0$ or $r_1 = 1$, respectively.

[4]The actual generated power typically differs from its set-point when generation controllers are sensitive to frequency deviations.

[5]We assume that a hybrid power plant is a source, thus $p_n^{\text{grid}} \geq 0$.

B. Restored Buses

We represent the fraction of energized buses via:

$$r_2 := \frac{1}{n_b}\sum_{s=1}^{n_b}\lambda\left(v_s\right) \text{ with } \lambda(x) := \begin{cases} 1 & \text{if } x \in [0.9, 1.1] \\ 0 & \text{otherwise}\,, \end{cases} \tag{23}$$

where $r_2 \in [0,1]$ and v_s is from (6). If $v_s \in [0.9, 1.1]$ for all s, then $r_2 = 1$; otherwise, $r_2 < 1$. This metric incentivizes a controller to maintain bus voltages within emergency levels by choosing an appropriate sequence of actions.

C. Restored Demand

We consider the fraction of the total power system demand that has been energized with:

$$r_3 := \frac{1}{n_\ell}\sum_{i=1}^{n_\ell}\min\left\{\frac{\ell_i}{\ell_i^{\text{target}}}, 1\right\}. \tag{24}$$

When each ℓ_i has reached or surpassed (because picked-up load is uncertain) its target, ℓ_i^{target}, then $r_3 = 1$. This metric serves to stimulate a controller to energize the demand which is the ultimate goal in a restoration process.

D. Transmission Congestion

We quantify the fraction of power transmission capacity that is available with respect to corresponding branch ratings via:

$$r_4 := 1 - \min\{p^\diamond, 1\} \text{ with } p^\diamond := \max_j\left\{\frac{|p_j^\dagger|}{p_j^{\dagger\text{rating}}}\right\} \tag{25}$$

where $r_4 \in [0,1]$. Here, $r_4 = 1$ signifies that the transmission system is unloaded. If $r_4 = 0$, it signifies that the transmission system has already congested. Controllers can be stimulated, for example, to re-dispatch generation to mitigate congestion.

E. Generation Strength

We model the strength of power generation (conventional or renewable) to pick up load and sustain it with:

$$r_5 := \max\left\{0, \frac{\sum_{n=1}^{n_g} p_n^{\text{firm}} - \sum_{i=1}^{n_\ell}\ell_i}{\sum_{n=1}^{n_g} p_n^{\text{firm}}}\right\} \tag{26}$$

where $r_5 \in [0,1]$. The n-th power plant can produce firm power, p_n^{firm}, if there is sufficient stored energy $e_n > e_n^{\min}$ (e.g., thermal or charged electrolyte) to sustain rated power generation, for at least τ_n hours (which can be defined by a grid operator) which implies:

$$p_n^{\text{firm}} = \begin{cases} p_n^{\max} & \text{if } \frac{e_n - e_n^{\min}}{p_n^{\max}} > \tau_n \\ 0 & \text{otherwise}\,. \end{cases} \tag{27}$$

F. Storage Level

The level of stored energy in the grid is sensed with:

$$r_6 := \frac{\sum_{n=2}^{n_g}\left(e_n - e_n^{\min}\right)}{\sum_{n=2}^{n_g}\left(e_n^{\max} - e_n^{\min}\right)} \tag{28}$$

where $r_6 \in [0,1]$. It serves to stimulate a controller to maintain relatively high reserves of stored energy if feasible.

G. Generation Load Sharing

During restoration, economic dispatch processes are typically offline. A judicious policy is that all generators are loaded in a similar percentage with respect to their capacities. We capture uneven generator loading with:

$$r_7 := 1 - \min\{\Delta\overline{p}, 1\} \tag{29}$$

$$\Delta\overline{p} := \max_n\left\{\left|\frac{p_n^{\text{grid}}}{p_n^{\max}} - \overline{p}^{\text{grid}}\right|\right\} \text{ with } \overline{p}^{\text{grid}} := \frac{1}{n_g}\sum_1^{n_g}\frac{p_n^{\text{grid}}}{p_n^{\max}} \tag{30}$$

where $\Delta\overline{p}$ models a deviation of percent loading among several generators. This metric is instrumental to stimulate restoration controllers to not overload particular generation assets.

H. Equity in Load Restoration

There are a variety of social objectives that can be considered in restoration. Here, we illustrate one of these with a metric of equity between load buses relative to a targeted amount of restoration. In particular, the restored demand as a percentage of its target value could significantly differ among buses. For example, the restored demand at one bus could be 30% whereas at another one could be 90%. That difference could suggest decision biases to serve particular loads even if a plan is technical. Equity in load restoration is sensed with:

$$r_8 := 1 - \min\{\Delta\overline{\ell}, 1\} \tag{31}$$

$$\Delta\overline{\ell} := \max_i\left\{\left|\frac{\ell_i}{\ell_i^{\text{target}}} - \overline{\ell}\right|\right\} \text{ with } \overline{\ell} := \frac{1}{n_\ell}\sum_1^{n_\ell}\frac{\ell_i}{\ell_i^{\text{target}}} \tag{32}$$

where $\Delta\overline{\ell}$ models a deviation of percent energized load from the average $\overline{\ell}$ among several loads. This metric is instrumental to stimulate restoration controllers to maintain a balance of the percent energized demand among several buses.

I. Restoration Performance

Grid operators select a restoration action based on multiple observations and objectives. An action can be decided based on the progress during the restoration of a weighted sum of the restoration performance metrics:

$$w \triangleq \boldsymbol{\alpha}^\top \boldsymbol{r} = \sum_{i=1}^{8}\alpha_i r_i. \tag{33}$$

The weights α_i satisfy $0 \le \alpha_i \le 1\ \forall i$ and $\sum_{i=1}^{8}\alpha_i = 1$. In the next section, we leverage a greedy controller using an optimal version of $\boldsymbol{\alpha}$. This is computed offline via a Bayesian optimization approach to maximize consumed energy. The indicator (33) is also used in the form of a *reward* to synthesize a state-feedback controller via deep reinforcement learning.

IV. State Feedback Controllers

We use greedy and reinforcement learning methods to steer the grid restoration process of Section II when performance is sensed with the metrics of Section III.

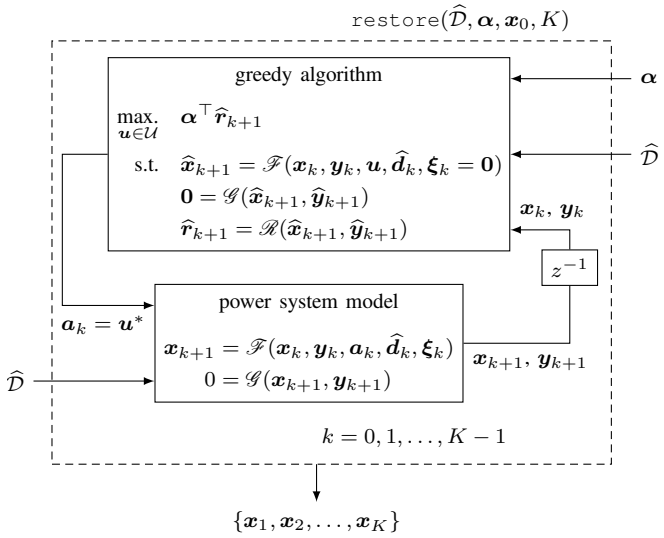


Fig. 4. Greedy restoration strategy of a power system from a blackout.

A. Greedy Restoration Technique

We formulate a greedy algorithm which autonomously outputs restoration actions using the observed state of the grid as depicted in Fig. 4. There, the pair $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ represents the present state of the power system. The algorithm estimates the next state of the grid $(\widehat{\boldsymbol{x}}_{k+1}, \widehat{\boldsymbol{y}}_{k+1})$ that could be attained by applying an action $\boldsymbol{u} \in \mathcal{U}$, hence leading to $\widehat{\boldsymbol{r}}_{k+1}$.

The control objective is to select a greedy action $\boldsymbol{u}^* \in \mathcal{U}$ of (4) that transfers $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ to a predicted $(\widehat{\boldsymbol{x}}_{k+1}, \widehat{\boldsymbol{y}}_{k+1})$ such that $\boldsymbol{\alpha}^\top \widehat{\boldsymbol{r}}_{k+1}$ is maximum. The controller is greedy because it only considers the next state when determining an action. The selected $\boldsymbol{u}^*$ is applied to the physical system which transfers the state of the grid to $(\boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1})$. We note that $(\boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1})$ is not necessarily equal to $(\widehat{\boldsymbol{x}}_{k+1}, \widehat{\boldsymbol{y}}_{k+1})$ because uncertainties cannot be captured in the controller. In the controller, we assume that $\boldsymbol{\xi}_k = \boldsymbol{0}$.

The selection of $\boldsymbol{\alpha}$ is a complex task analogous to tuning the parameters of a proportional controller. The objective is to select $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$ so that the energy served to the demand is maximum during a K-step restoration process. This implies posing the following optimization problem:

$$\begin{aligned} \max_{\boldsymbol{\alpha}} \quad & \sum_{k=1}^{K} \boldsymbol{\beta}^\top \boldsymbol{x}_k \\ \text{s.t.} \quad & \sum_{n=1}^{8} \alpha_n = 1 \text{ and } 0 \le \alpha_n \le 1,\ n \in \{1, 2, \ldots, 8\} \\ & \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_K\} = \texttt{restore}(\widehat{\mathcal{D}}, \boldsymbol{\alpha}, \boldsymbol{x}_0, K) \end{aligned} \tag{34}$$

where $\texttt{restore}(\widehat{\mathcal{D}}, \boldsymbol{\alpha}, \boldsymbol{x}_0, K)$ is illustrated in Fig. 4.[6] We clarify that the objective in (34) is the total energy served:

$$\sum_{k=1}^{K} \boldsymbol{\beta}^\top \boldsymbol{x}_k := \frac{1}{60} \sum_{k=1}^{K} \left(\sum_{i=1}^{n_\ell} \ell_i(t_k) \right) \tag{35}$$

in per unit hours which is extracted from each $\boldsymbol{x}_k$ ($k = 1, 2, \ldots, K$) by an appropriate selection of $\boldsymbol{\beta}$. The constant 1/60 appears in (35) because we choose $t_{k+1} - t_k = 1/60$ hours. Solving (34) is challenging because the problem is non-convex and the gradient information is unavailable. Note in (34) that the relation between the objective function and the decision vector is complex because a specific $\boldsymbol{\alpha}$ yields a particular $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_K\}$.

To solve (34), we resort to a zeroth-order optimization technique, viz., Bayesian optimization [29]. This method constructs Gaussian distributions along with confidence intervals based on interactions of the decision variables with the objective function. The maximum value of an acquisition function, which depends on the mean and variance of the distributions, creates the next data point to be evaluated. The acquisition function controls the trade-off between exploration and exploitation. The maximum of the lower-bound confidence interval dictates the search region as the upper bound must remain strictly larger than this value to prove a maximum of the objective function is reached. More technical details of the Bayesian optimization method are available in [29].

B. Deep Reinforcement Learning Approach

In contrast to the greedy method, deep reinforcement learning implicitly considers the outcome of the full future restoration process when steered from the present state. Here, the control actions are generated by an action-value function [30]:

$$\mathscr{Q}_{\mathcal{W}} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_{y'}} \times \mathcal{U} \times \mathbb{R}^{n_d} \mapsto \mathbb{R} \tag{36}$$

which is defined by a deep neural network, and where $\mathcal{W}$ represents a set of weights and biases or simply parameters. When $\mathcal{W} = \mathcal{W}^\star$, i.e., it is optimal, the function $\mathscr{Q}_{\mathcal{W}^\star}$ is used to output an approximation of the maximum discounted cumulative reward that a restoration process could achieve in an infinite-time horizon, i.e.:[7]

$$\max_{\boldsymbol{u}\in\mathcal{U}} \mathscr{Q}_{\mathcal{W}^\star}(\boldsymbol{x}_k, \boldsymbol{y}'_k, \boldsymbol{u}, \widehat{\boldsymbol{d}}_k) \approx \sum_{\iota=k}^{\infty} w_\iota \mu^{\iota-k}. \tag{37}$$

Here, $\boldsymbol{y}'_k = [p^{\text{grid}}_{1,k}; \boldsymbol{v}_k; \boldsymbol{p}^\dagger_k] \in \mathbb{R}^{n_{y'}}$ is extracted from $\boldsymbol{y}_k$ as specified in Section II-F, $w_\iota = \boldsymbol{\alpha}^{\star\top} \boldsymbol{r}_\iota$ is defined in (33) of Section III-I, and $\boldsymbol{\alpha}^\star$ is calculated in (34) of Section IV-A. We clarify that (37) quantifies a restoration process that is driven optimally from the present state $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ to a desired one $(\boldsymbol{x}^\star_K, \boldsymbol{y}^\star_K)$. In particular, the process is steered by a sequence of optimal actions $\{\boldsymbol{a}^\star_k, \boldsymbol{a}^\star_{k+1}, \ldots, \boldsymbol{a}^\star_{K-1}\}$ chosen from $\mathcal{U}$ while impacted by $\boldsymbol{d}_k, \boldsymbol{d}_{k+1}, \ldots \in \mathcal{D}$. Also, the discount factor $\mu \in (0, 1)$ stimulates early or delayed actions that will impact (37). For example, if $\mu \to 0$, the first instances of w_ι will dominate the right-hand side of (37).

[6] Alternatively, the coefficients α_i $\forall_i$ could be constrained as to reflect priorities on particular metrics. For example, transmission line congestion could be more important than generation strength so that $\alpha_4 > \alpha_5$.

[7] We use $\widehat{\boldsymbol{d}}_k$ because only the forecast is available for control synthesis.

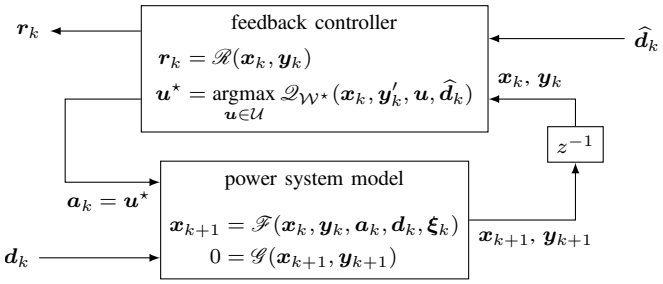


Fig. 5. Deployment of a reinforcement-learning agent for grid restoration.

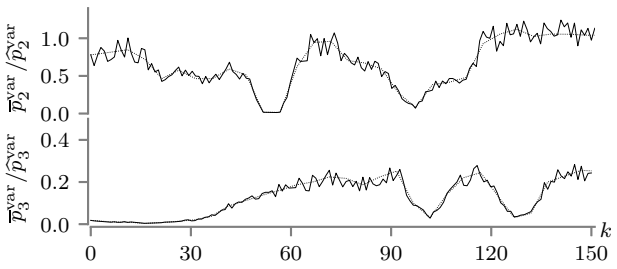


Fig. 6. Forecast (dotted lines) and actual (solid lines) solar power available at power plants CSP2 and S3 of Fig. 2 in per unit with base 100 MVA.

The application of $\mathscr{Q}_{\mathcal{W}^\star}$ for state-feedback control of the restoration of a power system is depicted in Fig. 5. The best control action, $\boldsymbol{a}_k$, is one $\boldsymbol{u} \in \mathcal{U}$ of (4) that maximizes $\mathscr{G}_{\mathcal{W}^\star}(\boldsymbol{x}_k, \boldsymbol{y}'_k, \boldsymbol{u}, \widehat{\boldsymbol{d}}_k)$. Here, the present state $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ and an estimate of exogenous input $\widehat{\boldsymbol{d}}_k$ are available information for control. We clarify that $\widehat{\boldsymbol{d}}_k \in \widehat{\mathcal{D}}$, which is a time-series data set, contains the forecast of solar irradiance, q.v. Section II. Notably, the action $\boldsymbol{a}_k$ will drive the grid to a next state $(\boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1})$; then, the calculation for the next action repeats.[8] The challenging task is the determination of $\mathcal{W}^\star$ in (37) and Fig. 5. The *training* process of $\mathscr{Q}_{\mathcal{W}}$ is explained with great detail in [30].

C. Mixed-Integer Linear Program

The MILP is not a feedback strategy, but is briefly presented for benchmarking. Here, a restoration plan is the solution to:

$$\begin{aligned} \max_{\boldsymbol{u}_k\in\mathcal{U}} \quad & \sum_{k=1}^{K} \boldsymbol{\beta}^\top \boldsymbol{x}_k \\ \text{s.t.} \quad & \boldsymbol{x}_{k+1} = \tilde{\mathscr{F}}(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{u}_k, \widehat{\boldsymbol{d}}_k)\ , k = 0, 1, \ldots, K-1 \\ & \boldsymbol{0} = \tilde{\mathscr{G}}(\boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1})\ , k = 0, 1, \ldots, K-1 \\ & \boldsymbol{x}\,, \boldsymbol{y} \in \mathcal{X}, \mathcal{Y} \end{aligned} \tag{38}$$

where $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ are given. The sets $\mathcal{X}$ and $\mathcal{Y}$ represent a variety of constraints that arise in a MILP formulation to constrain the state and output variables. The functions $\tilde{\mathscr{F}}$ and $\tilde{\mathscr{G}}$ are linearized forms of (1) and (2).

V. Case Study

We illustrate how to employ the model of Section II and the metrics in Section III as well as the feedback techniques in Section IV. These results are contrasted with a restoration plan from MILP in Section IV-C. We consider the modified version of the Western System Coordinating Council (WSCC) 9-bus power system that is depicted in Fig. 2.

All numerical studies were conducted in one core of a server with AMD 7742 CPUs running at a clock speed of 3.4 GHz. The power system of Section II is implemented in MATLAB 2021a [31] and MATPOWER 7.1 [32]. The greedy restoration technique in Section IV-A and reinforcement learning method of Section IV-B are implemented in Python 3.7.6 [33]. We also used the Adaptive Experimentation platforms (Ax) [29] to solve (34) with Bayesian optimization which yielded:

$$\boldsymbol{\alpha}^\star = [0.22, 0.001, 0.529, 0.001, 0.24, 0.001, 0.001, 0.007]^\top .$$

We note here that the weights emphasize r_1, r_3, and r_5 which makes sense because the objective of (35) is to energize the load. We also used Open AI Stable Baselines for pre-training, training and deployment of the controller in Section IV-B [34]. We used the following hyperparameters: $\eta = 0.99$, $\delta = 0.0005$, $\epsilon = 500$, which are the defaults in the tool. The MILP problem of Section IV-C is implemented in Julia 1.5.3 [35] and solved via Gurobi 9.1.0 [36]. Simulation parameters for model implementation of Section II are reported in Table I.

TABLE I
Simulation Parameters (per unit has base 100 MVA)

ℓ_1^{target}	1.3 p.u.	ℓ_2^{target}	0.9 p.u.	ℓ_3^{target}	1.0 p.u.
γ_i	0.9	γ_2	0.85	γ_3	0.95
$p_1^{\max}$	2.5	$p_2^{\max}$	1.6	$p_3^{\max}$	1.0
τ_1	10.0 h	τ_2	1.0 h	τ_3	1.0 h
η_n^{conv}	0.98 p.u.	η_n^{fwd}	0.95 p.u.	χ_n	$0.1p_n^{\max}$ p.u.
κ_i	0.1 p.u.	$\bar{\xi}_{\ell_i}$	0.2	$\bar{\xi}_{p_n}$	0.05
$\bar{\xi}_n^{\text{var}}$	0.2	$e_2^{\max}$	12.8 h	$e_2^{\min}$	1.28 h
$e_3^{\max}$	2.2 h	$e_3^{\min}$	0.22 h		

The ISO 6709 geographical locations of the solar data for G2 and G3 are: 40.95, -94.35 and 41.75, -93.45, respectively. The restoration trials for feedback control synthesis began at aleatory times of year 2006. To illustrate a specific case study, one-minute forecast instances of solar power for CSP2 and S3 from [37] are depicted in Fig. 6. For training of the GRA and DRL methods, the solar profiles in Fig. 6 are used to construct the sets $\mathcal{D}$ and $\widehat{\mathcal{D}}$, e.g., see Fig. 4. There, $k = 0$ corresponds to 10:00 am of January 1, 2006. The low energy limit for the storage of CSP2 and S3 is 10% of their energy maximum capacities, q.v. Fig. 2. At the start of restoration, we consider that the amount of stored energy at both power plants is 30% of their energy storage capacities.

Table II informs the value of the objective function (35) that is achieved by the greedy algorithm (GRA), deep reinforcement learning (DRL), and MILP strategies. It also displays the computational times for solving the respective control problem (training), sourcing a control action (inference), and generating a full restoration trail using the setup in Fig. 1 (simulation). Training times for GRA, DRL, and MILP pertain to determin-

[8]The grid can be stationary if the optimal action is do nothing.

TABLE II
OBJECTIVE FUNCTION & COMPUTATION TIMING.

	Objective (35)	Training	Inference	Simulation
GRA	650 MWh	3.35 min	0.1 s/action	13.0 s
DRL	660 MWh	15.56 min	0.64 ms/action	15.6 s
MILP	350 MWh (690 MWh)	21 s + 12.85 min	N/A	N/A

ing $\boldsymbol{\alpha}^*$ in Section IV-A, constructing $\mathscr{Q}_{\mathcal{W}^\star}$ in Section IV-B, and solving (38) in Section IV-C, respectively.

We learn from Table II that MILP achieves 390 MWh because the plan does not adapt to the uncertainties. MILP serves 690 MWh of energy when it is applied to a power system without uncertainties, q.v. Fig. 1. We also learn from Table II that synthesis of GRA controllers is faster than for DRL and MILP. On the other hand, the DRL controller can infer restoration actions up to one hundred times faster than the GRA one. The MILP calculation or training time of '21 s + 12.85 min' comprises of 21 s to determine a solution and 12.85 min to verify it. The computational complexity to scale GRA, DRL, and MILP to larger power systems respectively depends on the computational burden of: (i) the power flow calculations in (2), (ii) the back-propagation algorithm in [30], and (iii) the number of discrete variables and constraints in (38). Notably, inferring an action from MILP is not time consuming because the restoration plan is already computed. Although GRA and DRL need some CPU time to infer an action, they consider the present state and resource variability, thus they adapt to restoration challenges on the fly in contrast to MILP.

Figure 7 contrasts the restoration progress of the grid in Fig. 2 using GRA, DRL, and MILP. Figure 2 reveals that: (A) At $k = 3$, a 1.136 p.u. over-voltage occurs because a transmission line was energized and no load was connected. (B) At $k = 99$, there is a step change in the power transmitted by transformer T1 in Fig. 2 because the power sourced by S3, p_3^{grid}, drops from 0.71 to 0.03 p.u. Hence, p_1^{grid} of G1 increases to compensate the variability. (C) The power drop of S3 at $k = 99$ happens because its battery energy storage reaches its minimum level, i.e., $e_3^{\min} = 22$ MWh. (D) In contrast to GRA, the performance of DRL improves at $k = 73$ because battery charge of S3 rises. Hence, this generation resource becomes firm, q.v. (27). (E) The load ℓ_5, which is connected to bus B5 in Fig. 2, disconnects at $k = 23$. The reason is that the under-voltage protection activates because the bus voltage magnitude of B3 is 0.89 p.u., q.v. (14). (F) The performance of MILP is severely impacted at $k = 23$ because of the ℓ_5 disconnection.

Now we assess the performance of GRA and DRL by studying the values of w in Fig. 7 which are calculated with (33) at each point in time.

We found that the restoration process by DRL optimally charged/discharged the storage resources in contrast to the blackout recovery using GRA. This is because the control actions by GRA are greedy whereas DRL considers complete restoration trials. We recall via label (C) within Fig. 7 for GRA that e_3 (battery level of S3) reached its lowest admissible level 22 MWh at $k = 99$. Hence, p_3^{grid} is constrained by the availability of the solar resource $\overline{p}_3^{\text{var}}$, q.v. (20). The performance of MILP was challenged during deployment because a 20% uncertainty is applied to load pick-up, please see (14) and $\overline{\xi}_{\ell_i}$ in Table I. Uncertainties were not considered in the off-line optimization calculations.

Figure 8 illustrates the set of time-domain trajectories of restoration performance indices from Section III. We learn from Fig. 8 that: (A) Transmission congestion occurs at $k = 33$ because the branch connecting B1 to B4 is loaded to its limit as a result of the loading of G1, q.v. Fig. 2. (B) The battery depletion of S3 impacts branch congestion at $k = 99$. (C) The lowest unbalance of generation sharing is 0.33 at $k = 34$, q.v. (29). (D) The lowest level of load energization equity is 0.34 at $k = 28$, q.v. (31). (E) The system loads are fully restored when $k = 50$ and $k = 47$ when using GRA and DRL, respectively. (F) The MILP strategy did not fully restore the system loads because of the under-voltage event at $k = 23$ which challenged the pre-computed plan. (G) Generation strength when using MILP increases because of the disconnection of the load ℓ_5 at $k = 23$.

VI. CONCLUSION

This paper has reported a dynamic model of the restoration process of a power system which is used to synthesize decision-making controllers for grid restoration. We designed two state-feedback controllers that leverage the theory of greedy and deep reinforcement learning algorithms. The design process was driven by a set of eight restoration performance metrics which quantify the restoration of a grid in a relatively low dimensional space. An advantage of the state feedback controllers is that decisions are made based on the present state of the grid, which can account for the variability of renewable resources. We found that the DRL controller is roughly one hundred times faster than the GRA one; however, training time for DRL is nearly five times longer than identifying the coefficients for GRA using Bayesian optimization. Future work will address larger power systems, the incorporation of more uncertainties such as breaker failures, and simulating the restoration process using electromagnetic transient programs.

VII. ACKNOWLEDGMENTS

This work was authored in part by the National Renewable Energy Laboratory (NREL), operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development Program at NREL. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. Funding was provided to Iowa State University through awards SUB202010304 and SUB202110418.

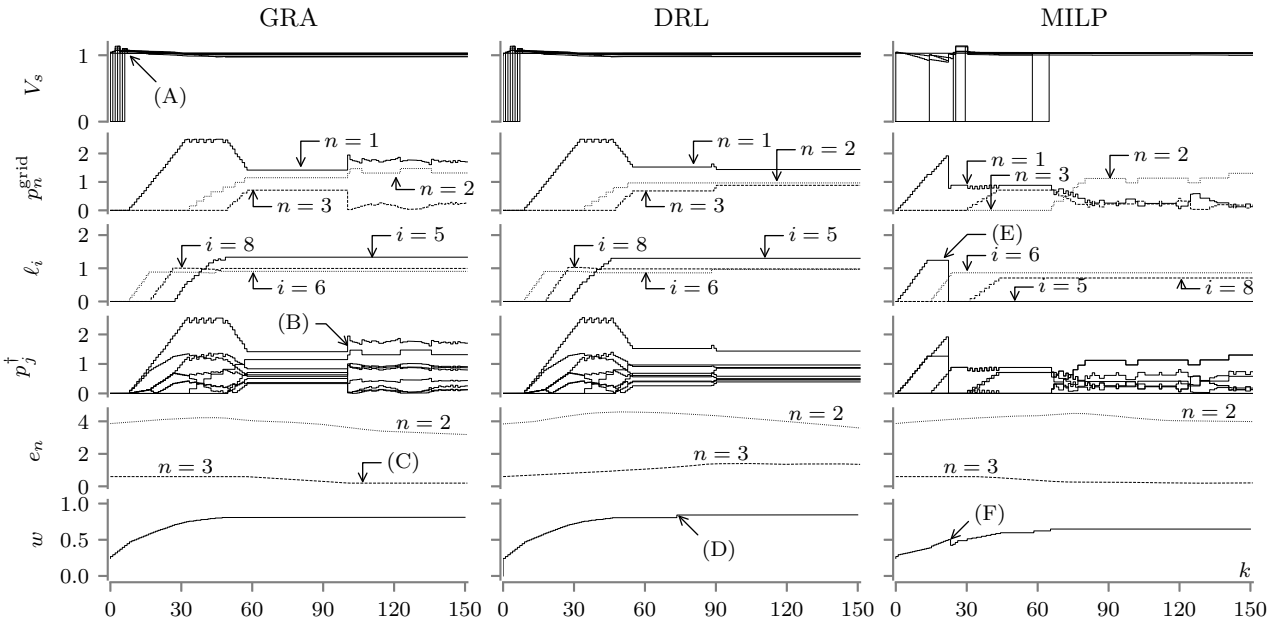


Fig. 7. Time-domain restoration trajectories for the modified WSCC power grid in Fig. 2. Quantities are in per unit with base 100 MVA.

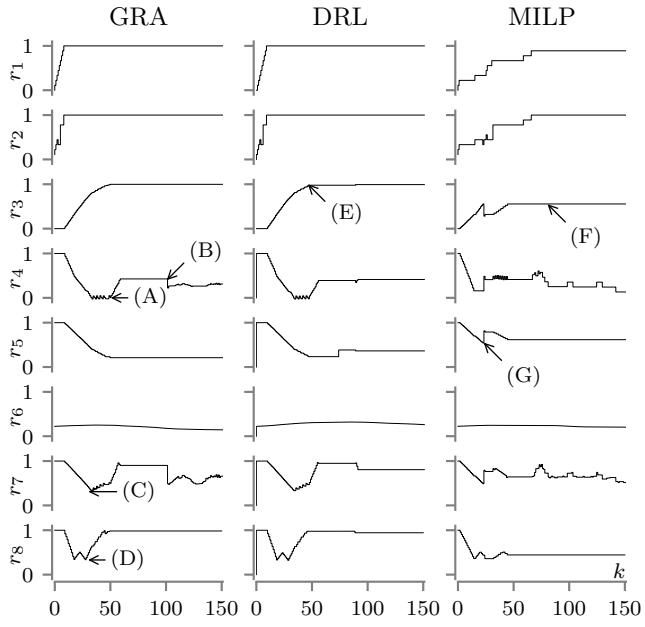


Fig. 8. Time-domain restoration performance metrics.

REFERENCES

[1] M. Adibi, P. Clelland, L. Fink, H. Happ, R. Kafka, J. Raine, D. Scheurer, and F. Trefny, "Power system restoration - a task force report," *IEEE Power Eng. Rev.*, vol. 7, no. 5, May 1987.

[2] J. Giri, D. S. Kirschen, S. M. Shahidehpour, and J. Zaborszky, "New approaches in power system restoration," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1428–1434, Nov. 1992.

[3] W. I. Rowen, "Operating characteristics of heavy–duty gas turbines in utility service," *The Amer. Soc. Mechanical Engineers*, vol. 5, pp. 1–13, Jun. 1988.

[4] J. R. Gracia *et al.*, "Hydropower plants as black start resources," Oak Ridge National Laboratory, Tech. Rep. ORNL/SPR-2018/1077, May 2019.

[5] "Black system South Australia 28 September 2016 - final report," Australian Energy Market Operator (AEMO), Tech. Rep., Mar. 2017.

[6] W. Jang, H. Huang, K. R. Davis, and T. J. Overbye, "Considerations in the automatic development of electric grid restoration plans," in *Proc. 52nd North Amer. Power Symp.*, Tempe, AZ, USA, Apr. 11–13 2020.

[7] G. Patsakis, D. Rajan, I. Aravena, J. Rios, and S. Oren, "Optimal black start allocation for power system restoration," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6766–6776, Nov. 2018.

[8] D. Velasco de la Fuente, C. L. Trujillo Rodríguez, G. Garcerá, E. Figueres, and R. Ortega Gonzalez, "Photovoltaic power system with battery backup with grid-connection and islanded operation capabilities," *IEEE Trans. Ind. Electron.*, vol. 60, no. 4, pp. 1571–1581, Apr. 2012.

[9] M. D. Tabone and D. S. Callaway, "Modeling variability and uncertainty of photovoltaic generation: A hidden state spatial statistical approach," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2965–2973, Dec. 2015.

[10] A. Golshani, W. Sun, Q. Zhou, Q. P. Zheng, and Y. Hou, "Incorporating wind energy in power system restoration planning," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 16–28, Jan. 2019.

[11] R. J. Kafka, D. R. Penders, S. H. Bouchey, and M. M. Adibi, "Role of interactive and control computers in the development of a system restoration plan," *IEEE Trans. Power App. Syst.*, vol. PAS-101, no. 1, pp. 43–52, Jan. 1982.

[12] J. Li, H. You, J. Qi, M. Kong, S. Zhang, and H. Zhang, "Stratified optimization strategy used for restoration with photovoltaic-battery energy storage systems as black-start resources," *IEEE Access*, no. 7, pp. 127 339–127 352, Aug. 2019.

[13] A. S. Bretas and A. G. Phadke, "Artificial neural networks in power system restoration," *IEEE Trans. Power Del.*, vol. 18, no. 4, pp. 1181–1186, Oct. 2003.

[14] D. Ye, M. Zhang, and D. Sutanto, "A hybrid multiagent framework with Q-learning for power grid systems restoration," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2434–2441, Nov. 2011.

[15] S. Das, S. Bose, S. Pal, N. N. Schulz, C. M. Scoglio, and B. Natarajan, "Dynamic reconfiguration of shipboard power systems using reinforcement learning," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 669–676, May 2013.

[16] L. R. Ferreira, A. R. Aoki, and G. Lambert-Torres, "A reinforcement learning approach to solve service restoration and load management simultaneously for distribution networks," *IEEE Access*, vol. 7, pp. 145 978–145 987, Oct. 2019.

[17] J. Wu, B. Fang, J. Fang, X. Chen, and C. K. Tse, "Sequential topology recovery of complex power systems based on reinforcement learning,"

Physica A: Statistical Mechanics and its Applications, vol. 535, no. 122487, Dec. 2019.
[18] M. Glavic, "(Deep) reinforcement learning for electric power system control and related problems: A short review and perspectives," *Annu. Reviews Control*, vol. 50, pp. 22–35, Jul. 2019.
[19] L. Duchesne, E. Karangelos, and L. Wehenkel, "Recent developments in machine learning for energy systems reliability management," *Proc. IEEE*, vol. 108, no. 9, pp. 1656–1676, May 2020.
[20] B. Chiu *et al.*, "Resilience framework, methods, and metrics for the electricity sector," ITSLC, Tech. Rep., Oct. 2020.
[21] G. Porro *et al.*, "Grid modernization: Metrics analysis (GMLC1.1) selected literature review and mapping," GMLC, Tech. Rep., Apr. 2020.
[22] S. Abourida *et al.*, "Real-time power system simulation: EMT vs. phasor," OPAL-RT, Tech. Rep. opWP150620-sa-revA, 2016.
[23] WECC Renewable Energy Modeling Task Force, "WECC PV power plant dynamic modeling guide," WECC, Tech. Rep., Apr. 2014.
[24] *Wind turbines - Part 27-1: Electrical simulation models - Wind turbines*, International Electrotechnical Commission Std. IEC 61 400-27-1, Rev. 1.0, Feb. 2015.
[25] A. Honrubia-Escribano *et al.*, "Field validation of a standard type 3 wind turbine model for power system stability, according to the requirements imposed by IEC 61400-27-1," *IEEE Trans. Energy Convers.*, vol. 33, no. 1, pp. 137–145, Mar. 2018.
[26] Z. Stanislaw, *Systems and Control*, ser. The Oxford Series in Electrical and Computer Engineering, A. S. Sedra, Ed. New York, NY, USA: Oxford Univ. Press, 2003.
[27] *System Restoration from Blackstart Resources*, North American Electric Reliability Corporation Std. EOP-005-3, Apr. 2019.
[28] *Generator Frequency and Voltage Protective Relay Settings*, North American Electric Reliability Corporation Std. PRC-024-3, Apr. 2019.
[29] Facebook Open Source, "Ax: Bayesian optimization," Accessed: Aug. 7, 2021. [Online]. Available: https://ax.dev/docs/bayesopt.html
[30] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015. [Online]. Available: http://dx.doi.org/10.1038/nature14236
[31] *MATLAB R2021a*, The MathWorks Inc., Natick, MA, USA, 2021.
[32] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Jun. 2011.
[33] Python Software Foundation, "Python 3.7," Accessed: Aug. 10, 2021. [Online]. Available: https://www.python.org/
[34] "Stable baselines," 2018–2021, Accessed: Aug. 7, 2021. [Online]. Available: https://stable-baselines.readthedocs.io/en/master/guide/quickstart.html.
[35] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM review*, vol. 59, no. 1, pp. 65–98, Feb. 2017.
[36] Gurobi Optimization, LLC, *Gurobi Optimizer Reference Manual*, Accessed: Mar. 3, 2022. [Online]. Available: https://www.gurobi.com
[37] National Renewable Energy Laboratory, "Solar power data for integration studies," Accessed: Mar. 3, 2022. [Online]. Available: https://www.nrel.gov/grid/solar-power-data.html