Quantile Fourier regressions for decision making under uncertainty

Arash Khojaste[1], Geoffrey Pritchard[2*], Golbon Zakeri[1]

[1]Dept. of Mechanical and Industrial Engineering, University of Massachusetts – Amherst, Amherst, 01002, MA, United States.
[2*]Dept. of Statistics, University of Auckland, Private Bag 92019, Auckland, 1142, New Zealand.

*Corresponding author(s). E-mail(s): g.pritchard@auckland.ac.nz;
Contributing authors: akhojaste@umass.edu; gzakeri@umass.edu;

Abstract

We consider Markov decision processes arising from a Markov model of an underlying natural phenomenon. Such phenomena are usually periodic (*e.g.* annual) in time, and so the Markov processes modelling them must be time-inhomogeneous, with cyclostationary rather than stationary behaviour. We describe a technique for constructing such processes that allows for periodic variations both in the values taken by the process and in the serial dependence structure. We include two illustrative numerical examples: a hydropower scheduling problem and a model of offshore wind power integration.

Keywords: Markov decision processes, quantile regression, offshore windpower, hydropower scheduling

1 Introduction

We consider stochastic optimization problems in which the underlying randomness takes the form of a discrete-time univariate stochastic process with periodically varying behaviour. Our approach will be to approximate the underlying stochastic process by a finite-state Markov chain in which both the interpretation of the states and the Markov transition probabilities are periodically varying (time-inhomogeneous). The optimization problem is then modelled as a Markov decision process.

Many real-world planning problems involve decisions of two kinds, which we characterize as "investment" and "operations". In the investment phase, decisions are taken which create or preclude possible courses of action in the later operations phase. Examples include capacity planning for electric power, management and operations in forestry, and literal financial investments. The early decisions (*e.g.* on the construction or retirement of power-generation plants) can have a large effect on the cost and robustness of later operations (*e.g.* efficient and reliable dispatch of electricity). Uncertainty plays a key role: the investment decisions must be made without full knowledge of the operational environment in which their consequences will play out.

The operations phase typically represents a long-term future, and so it is often further subdivided into many stages. At each stage, some of the uncertainty is resolved, and recourse actions become possible in response to the new information. This is the paradigm of multi-stage stochastic programming (see *e.g.* [1, 4, 5, 20]). The uncertainty in such problems can be represented by scenario trees that branch at each operational stage. For example, investment decisions for electricity generation would be followed by a multitude of hour-long operational stages (8760 of them in each year, perhaps for many future years), with multiple random quantities realized at each stage. It is evident that due to the exponential growth in the number of scenarios and the distant time horizon, such problems can become prohibitively large. Much research effort has been spent on developing decomposition approaches and approximation schemes that iteratively solve a limited version of the problem in pursuit of a solution (see *e.g.* [1, 14, 25]).

This paper takes a different approach: we propose to represent the operations phase via the stationary behaviour of a discrete Markov decision process (MDP). While the usual exposition of MDPs posits a single endlessly repeated problem (given that the system is in one of finitely many possible states, choose one of finitely many possible actions which will determine the probability distribution of the state in the next incarnation), it is readily adaptable to models in which the states, actions, and associated parameters vary in a periodic cycle. The solution is then a cyclostationary rather than a stationary process.

This succinct representation of uncertainty can then be embedded within the investment phase of the problem, allowing the effects of investment decision-making on subsequent operations to be easily evaluated. The present paper, however, will confine its focus to the modelling and solution of the operations phase. We will outline a systematic approach for constructing a suitable Markov decision process via quantile Fourier regression. Quantile regressions are ideal for modelling continuously and periodically varying random phenomena ([18]) in which the probability distributions must be different at each point in the periodic cycle (*e.g.* on each day of the year). We also describe an apparently new method of fitting time-inhomogeneous Markov transition probabilities, so that not only the random values themselves, but also their serial dependence structure, can be subject to periodic variation.

We illustrate the techniques developed in this paper with two applications. In section 5 we discuss a hydro-thermal power scheduling problem with seasonal variation in inflows. In section 6 we consider a more ambitious model of offshore windpower integration subject to both diurnal and annual variations.

2 Quantile Fourier regression

A discrete-time stochastic process $(X_t)_{t=0}^{\infty}$ is *cyclostationary* if, for any integers $m_1, \ldots, m_k$, the joint probability distribution of $X_{t+m_1}, \ldots, X_{t+m_k}$ is a periodic function of t. For the purposes of the present paper, the period p will always be an integer; thus

$$(X_{t+m_1}, \ldots, X_{t+m_k}) \stackrel{D}{=} (X_{t+m_1+p}, \ldots, X_{t+m_k+p}).$$

For further and more general discussion of cyclostationary processes, the reader is referred to [10].

In modelling some phenomenon as a cyclostationary process, we might perhaps begin by describing the probability distribution of X_t as a function of t. We shall try to avoid making specific assumptions about the shape(s) of this distribution. However, if the period p is a large integer—*e.g.* a process with hourly or daily time steps and annual periodicity—it is desirable to assume some structure in the way the distribution varies as a periodic function of t, rather than attempting to fit p completely unrelated distributions. (The most serious objection to the latter approach is that each such distribution must be estimated from a subsample comprising only $\frac{1}{p}$ of the available data; this subsample will usually be too small to do the job well.) We also assume that p is known.

One approach which achieves these aims is quantile Fourier regression. For a fixed $\tau \in (0, 1)$, the τ-quantile of the distribution of X_t is modelled as a function $q_\tau(t)$ with period p drawn from a finite-dimensional space of such functions:

$$P(X_t \leq q_\tau(t)) = \tau \qquad \text{where} \qquad q_\tau(t) = \sum_{j=1}^{d} \beta_j b_j(t). \tag{1}$$

Here $\beta_1, \ldots, \beta_d$ are coefficients to be fitted from data, while $b_1(t), \ldots, b_d(t)$ are a basis for the chosen function space. Suitable function spaces include periodic splines and wavelets ([2]). But in the present paper, we make the same choice of basis as Jean-Baptiste Joseph Fourier: $d = 2r + 1$ is odd and

$$b_0(t) = 1 \quad \text{and} \quad b_{2k-1}(t) = \cos(k\omega t), \quad b_{2k}(t) = \sin(k\omega t), \text{ for } k = 1, \ldots, r, \tag{2}$$

where $\omega = 2\pi/p$. For small values of r, the quantile is thus constrained to a smooth and fairly simple variation with t. After constructing such models for several different values of τ, we have a (partial) description of the distribution of X_t throughout the periodic cycle. Estimation of the coefficients β_j is a well-understood technique ([15]) which reduces to linear programming, and so is computationally straightforward.

The simplest special case of (1) occurs when the $b_j(t)$ are piecewise constant. This occurs, for example, when there is annual periodicity and the stationary distribution of X_t is assumed to be constant within each month of the year, so that there are 12 (unrelated) distributions to fit. With this approach, estimation could hardly be simpler: to estimate, say, the 75th percentile of the distribution for January, we need only calculate the 75th percentile of all the data points observed in January of any year. However, this is a poor choice of functional form when our ultimate intention

is to incorporate the resulting model into a stochastic optimization problem. The discontinuous changes in the distribution at certain points in the periodic cycle are likely to distort optimal decisions near those points (a "month-end effect"). It is for this reason that the present paper considers only continuous models for the quantiles.

It is possible that there is more than one periodic cycle. For example, a process with hourly time steps may have both diurnal and annual periodicity. For simplicity, we assume in this paper that the periods are commensurate, so that there is a single overall period which can be expressed as an integer multiple of each individual period. The modelling problem then reduces to finding a function space for the quantiles which incorporates all of the periodic variations present.

3 A Markov model of serial dependence

Describing the probability distribution of X_t for each t does not complete the modelling task; we should also consider the serial dependence structure of the process.

Perhaps the simplest model of serial dependence is the finite-state Markov chain. With the pointwise stationary distribution(s) of our process modelled via their quantiles as in section 2, we are well-placed to adopt such a model. The quantile representation naturally lends itself to the discretization of a univariate process to a finite-state one: if we have functions modelling quantiles $\tau_1 < \cdots < \tau_{m-1}$, then for $i = 2, \ldots, m-1$ the interval $[q_{\tau_{i-1}}(t), q_{\tau_i}(t)]$ can be represented by state i. We also have extreme intervals $(-\infty, q_{\tau_1}(t))$ and $[q_{\tau_{m-1}}(t), \infty)$ represented by states 1 and m respectively. For convenience in referring to these, we define $\tau_0 = 0$, $\tau_m = 1$, $q_0(t) = -\infty$, and $q_1(t) = \infty$.

Inference of the Markov transition probabilities can be carried out in several ways. The most straightforward is to fit a single transition matrix to all of the available data: if there are n_{ij} transitions observed from state i to state j, then the corresponding element of the fitted transition matrix is $p_{ij} = n_{ij} / \sum_{k=0}^{m} n_{ik}$. A small refinement is to observe that for a Markov chain constructed in this way, the stationary probability distribution is already known: the stationary probability π_i of state i should be $\tau_i - \tau_{i-1}$ $(i = 1, \ldots, m)$. In the special case of equally spaced quantiles ($\tau_i = i/m$), the stationary distribution is uniform ($\pi_i = 1/m$ for each i); equivalently, the transition matrix is doubly stochastic. A transition matrix constrained to satisfy this requirement can be estimated *e.g.* by the Sinkhorn-Knopp algorithm ([21, 22]). Such a model is described in detail in [19].

More ambitiously, we can consider that the serial dependence structure of the original process may also exhibit variation throughout the periodic cycle. To accommodate such variation, the transition probabilities $p_{ij}(t)$ should be allowed to be time-inhomogeneous: periodic functions of t, with the same period p as the original process. That is,

$$p_{ij}(t) = \sum_{\ell=0}^{q} \gamma_{ij\ell} b_\ell(t) \tag{3}$$

where $b_0(t), \ldots, b_q(t)$ are periodic basis functions and $\gamma_{ij\ell}$ are fitted coefficients.

This is a separate and different kind of periodicity from that contemplated in Section 2: while (1) describes a periodic variation in the absolute meaning of the

Markov states, (3) describes a periodic variation in the transition behaviour between states. Suppose, for example, that the original process (X_t) represents daily rainfall; then the highest Markov state m represents a rainy day. For any time of year t, the function $q_{\tau_{m-1}}(t)$ gives information on the likely amount of rain falling on one rainy day, while $p_{mm}(t)$ gives information on the likelihood of a sequence of consecutive rainy days. (The length of any such sequence, if short relative to the period p, has approximately a geometric distribution with parameter $p_{mm}(t)$.)

The periodic function space used in (3) can be the same as that used in (1), or different. For the examples in the present paper, we will again use the Fourier basis.

Estimation of the coefficients $\gamma_{ij\ell}$ can be done by a maximum likelihood procedure. If we have observed transitions at times t_k from state i_k to state j_k, for $k = 1, \ldots, n$, then the likelihood is $L = \prod_{k=1}^{n} p_{i_k j_k}(t_k)$, and so the coefficients should be chosen to minimize

$$-\log L = -\sum_{k=1}^{n} \log \left(\sum_{\ell=0}^{q} \gamma_{i_k j_k \ell} b_\ell(t_k) \right)$$

subject to the constraints

$$\begin{aligned}
0 \le \sum_{\ell=0}^{q} \gamma_{ij\ell} b_\ell(t) \le 1 \qquad & \text{for all } i,\ j, \text{ and } t \\
\sum_{j=1}^{m} \sum_{\ell=0}^{q} \gamma_{ij\ell} b_\ell(t) = 1 \qquad & \text{for all } i \text{ and } t \\
\sum_{i=1}^{m} \pi_i \sum_{\ell=0}^{q} \gamma_{ij\ell} b_\ell(t) = \pi_j \qquad & \text{for all } j \text{ and } t
\end{aligned}$$

where π_i is the known stationary probability of state i noted above. This is a nonlinear convex optimization problem. Note the requirement that the constraints hold for all times t, rather than for only those times t_k at which transitions have been observed. In the case of the Fourier basis (2), the equality constraints reduce to

$$\sum_{j=1}^{m} \gamma_{ij0} = 1 \quad \text{for all } i \qquad \text{and} \qquad \sum_{j=1}^{m} \gamma_{ij\ell} = 0 \quad \text{for all } i \text{ and } \ell > 0$$

$$\sum_{i=1}^{m} \pi_i \gamma_{ij0} = \pi_j \quad \text{for all } j \qquad \text{and} \qquad \sum_{i=1}^{m} \pi_i \gamma_{ij\ell} = 0 \quad \text{for all } j \text{ and } \ell > 0.$$

4 A Markov decision process

Equipped with a Markov process that describes an underlying natural phenomenon, *e.g.* rainfall, we can proceed to the decision making part of the model. Markov decision processes span a wide range of models, that include finite or infinite state as well as discrete or continuous processes. For the purposes of this paper, we focus on finite discrete MDPs.

For the governance of our system, we may expand the state space to include additional states beyond the underlying natural phenomena. For instance, energy-related problems may include the levels of charge stored in a battery or other form of storage, in addition to rainfall and demand states. We will assume that at any time t and in any given state i, we can take a range of actions (or decisions) $k \in \{0, 1, \cdots, K\}$. The effect of taking an action is twofold: firstly, it incurs a cost; secondly, it determines the probability distribution of the state at time $t + 1$.

Markov decision processes select policies that lead to a desired objective, such as minimizing long-run expected average cost per unit time. A randomized policy will present the decision maker with a probability distribution $(d_{i1t}, d_{i2t}, \cdots, d_{iKt})$ for the action k to be chosen when in state i at time t. Given such a policy, the system can operate as a cyclostationary process in the sense given in Section 2; this represents its long-run behaviour under the specified policy. We can find the policies minimizing the average (over the periodic cycle) expected cost by forming the following linear program. Note that choosing an action affects the probability of reaching other states as well as the immediate cost accrued, both of which will have an impact on the steady state distribution of cost, hence the expectation of cost.

$$\begin{array}{lll} \text{Minimize} & \sum_t \sum_{i=0}^{M} \sum_{k=0}^{K} C_{ikt} y_{ikt} & \\ \text{s/t} & \sum_{i=0}^{M} \sum_{k=0}^{K} y_{ikt} = 1 & \forall t \\ & \sum_{k=0}^{K} y_{j,k,t+1} - \sum_{i=0}^{M} \sum_{k=0}^{K} y_{ikt} P_{ijt}(k) = 0 & \forall t, \forall j \\ & y_{ikt} \geq 0 & \forall i, \forall k, \forall t. \end{array} \tag{4}$$

This linear program is an extended version of those given in [11, 24], where we have introduced periodic time inhomogeneity, hence the dependence on t. The parameter $P_{ijt}(k)$ represents the probability of transition from state i to state j, when action k is taken at time t. The parameter C_{ikt} is the cost incurred by being in state i at time t and taking action k. Decision variable y_{ikt} represents the cyclostationary probability that the system is in state i and action k is undertaken at time t. Once the optimal y_{ikt} variables are obtained, we can determine d_{ikt} using $d_{ikt} = \frac{y_{ikt}}{\sum_{k=0}^{K} y_{ikt}}$.

5 Application: reservoir management

In this section, we apply our technique to a hydro-power reservoir management problem of the kind described in [23]. The underlying data are a univariate time series of weekly natural inflows to the Waitaki River system in New Zealand; see [18] for more details on this data set. Although the real system comprises multiple catchments, reservoirs, and power generation sites, we treat it here as a single equivalent-energy reservoir, with inflows expressed in power units (megawatts).

Figure 1 (left panel) shows the data set together with quantile regression models for the annual variation of the 10th, 50th, and 90th percentiles. These models were constructed with the Fourier basis (2) with $r = 2$; that is, the fitted curves are second-order trigonometric polynomials.

These three quantile models were then used to model the serial dependence structure as a four-state Markov chain as described in Section 3. All sixteen transition probabilities were permitted to vary annually as simple sinusoids, by using the Fourier

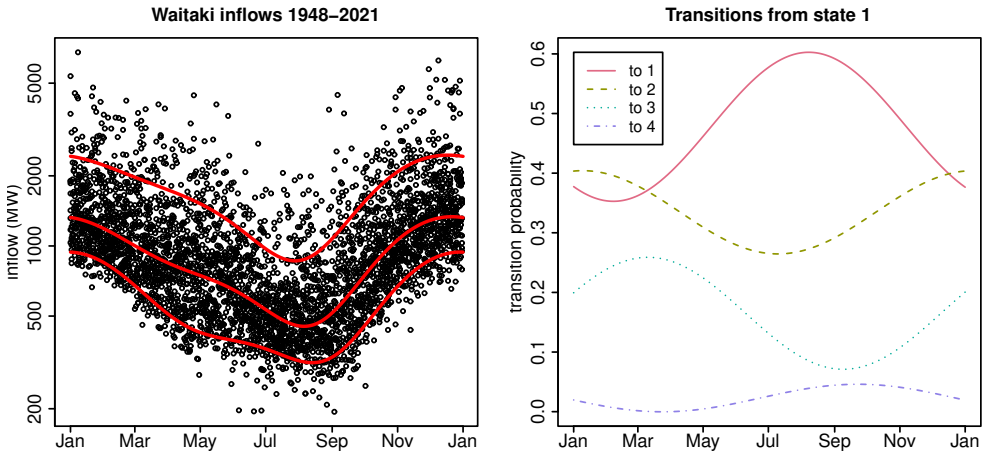

Fig. 1 The Waitaki River model. Left panel: weekly inflow data, with fitted 10th, 50th, and 90th percentile models. Right panel: fitted transition probabilities from the first (lowest-flow) Markov state.

basis (2) with $r = 1$. The fitted probabilities for transitions from state 1 (lowest flow) to other states are shown in Figure 1 (right panel).

In addition to the state of the inflow process, the state space of the full reservoir-management problem needs to describe the storage state of the reservoir. The energy storage capacity of the real Waitaki River system is approximately 2500 gigawatt-hours (GWh) [7], but we will here consider an illustrative problem in which the reservoir capacity is only 840 GWh. (Downsizing the reservoir is done principally to make the problem more interesting, by increasing the probabilities of shortages and overflows.) The Markov decision process requires discrete states, which we create by discretizing the stored energy into blocks of 16.8 GWh. Since the time step is one week, this block size can be conveniently expressed in power units as 100 MW. The reservoir capacity is 50 blocks, and so there are 51 possible storage states. With the four possible inflow states, the size of the state space is 204.

The Markov inflow process must now be further developed to deliver energy inflows in blocks of the same size as used for the reservoir storage. Each state i of the original process discussed in Section 3 corresponds only to a time-varying interval $[q_{\tau_{i-1}}(t), q_{\tau_i}(t)]$ containing the amount of energy inflow. Our chosen block size is small enough that this interval usually contains least two different multiples of the block size. So, we can take the inflow amount to be one of these multiples, chosen at random independently of the underlying Markov process. Note that this approach preserves the continuous flavour of the original inflow model: even though the inflow amounts are now discrete, their probabilities can be thought of as varying continuously with time t.

We determine the distribution of the inflow, conditional on state i and time t, in the following way. For states other than the highest and lowest, the original inflow

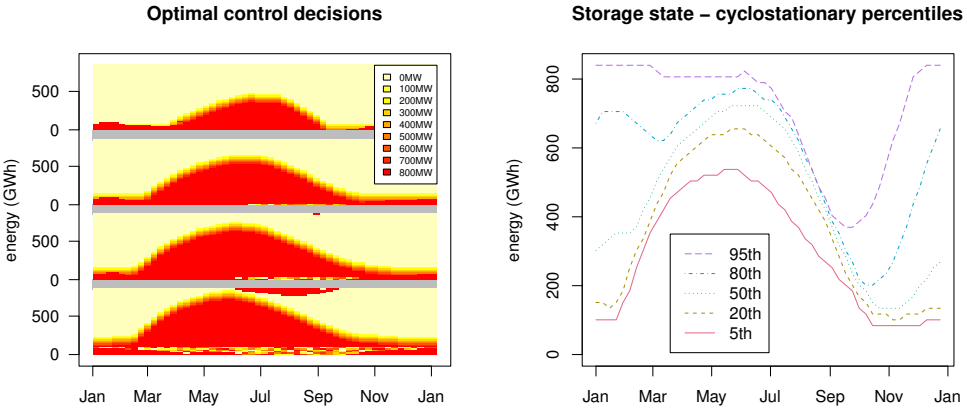

Fig. 2 The Waitaki River MDP solution. Left panel: the optimal decisions (non-hydro generation quantities). Right panel: the cyclostationary distribution of stored energy as a function of time.

data corresponding to state i are normalized by the transformation

$$x \mapsto \frac{x - q_{\tau_{i-1}}(t_x)}{q_{\tau_i}(t_x) - q_{\tau_{i-1}}(t_x)}$$

(where t_x is the time at which inflow x was observed), giving an empirical probability distribution on $[0, 1]$. The multiples of the block size permissible at time t are then transformed in the same way, and assigned probabilities to match the empirical distribution as closely as possible. For the lowest and highest states, the corresponding intervals are semi-infinite: $(-\infty, q_{\tau_1}(t)]$ and $[q_{\tau_{m-1}}(t), \infty)$. In these cases, the normalization transformation used is a simple division by the finite endpoint.

For our example reservoir management problem, we require the system to meet a constant demand for 1400 MW of electric power. As this exceeds the mean hydroenergy inflow, we also include 800 MW of dispatchable non-hydro power generation capacity, with an assumed cost of \$50 per megawatt-hour (MWh). Any unserved demand incurs a cost of \$1000/MWh. The hydropower generation capacity is taken to be 1500 MW.

We seek a cyclostationary solution with annual period, consisting of 52×7-day time steps, for which the expected cost per annum of non-hydro power and unserved demand is minimized. At each time step, the available actions consist of dispatching some amount of non-hydro generation; since this must also be delivered in 100 MW blocks, there are 9 different actions available. Thus, the variables y_{ikt} in (4) number $204 \times 9 \times 52 = 95472$.

Figure 2 illustrates the solution of the MDP. Relatively abundant summer inflows create a wide range of possible storage states at the beginning of the year. But by the beginning of May (late autumn), the optimal control has used non-hydro generation

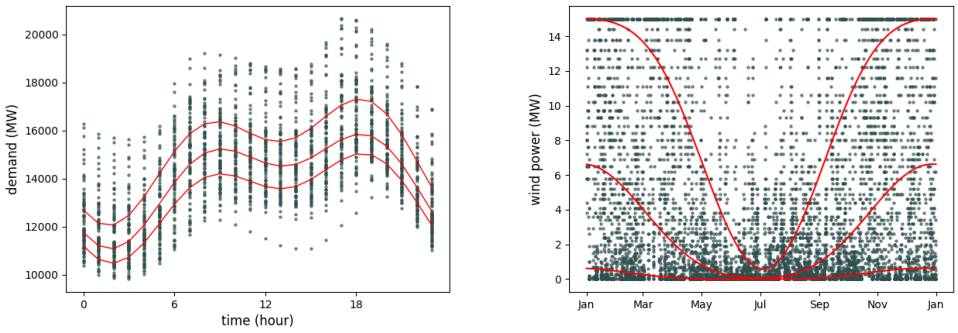

Fig. 3 Demand and wind power processes, with 25th, 50th, and 75th percentiles estimated by quantile Fourier regression with $r = 2$ and superimposed on 2018 data. Left panel: diurnal variation in New England winter demand. Right panel: annual variation in generation by one IEA 15 MW offshore wind turbine.

to ensure that the storage reservoir is fairly full, so that the stored energy can be drawn upon during the low-inflow winter period.

6 Application: thermal backup of offshore wind

With the increasing penetration of renewable sources of electricity generation into the world's electricity systems, it is imperative to have sufficient backup (or firming) for intermittent renewables and to operate it efficiently to avoid costly curtailing of demand for electricity. In the United States there are a number of mandates to procure electricity generation from offshore wind; for instance, Massachusetts has set a target of 3.2 gigawatts (GW) of offshore wind capacity by 2035 [16].

Our first example concerns the operation of a system of backups to cover any shortfall in power generated from offshore wind resources. The model's underlying univariate stochastic process represents the New England regional demand for electric power, net of modelled offshore wind power generation.

We use hourly electricity demand data for the ISO-New England grid from 2006–2020 as reported to the Federal Energy Regulatory Commission (FERC) [8]. Figure 3 (left panel) illustrates the diurnal variation, showing a double-peaked pattern typical of winter (but not summer) electricity consumption. This pattern indicates that quantile modelling with annual period must also capture daily periodicity to at least the second harmonic and that the shape of the diurnal variation should be modulated by the time of year. That is, the basis functions in (1) should include *e.g.* $\cos(\omega_1 t)\cos(2\omega_2 t)$ and similar functions, where $\omega_1 = 2\pi/\text{year}$ and $\omega_2 = 2\pi/\text{day}$.

For wind power, the underlying data are hourly wind speeds recorded by the National Oceanic and Atmospheric Administration (NOAA) National Data Buoy Center [17]. As a reference station for evaluating offshore wind power generation, we used Buoy 44025, in the New York Bight area. We chose the IEA 15 MW reference turbine

[9] as our power generation mechanism, using the power curve and other specifications (hub height, etc.) for this turbine to make the wind-speed to power conversion. Details of the power curve modeling can be found in [3, 13]. Figure 3 (right panel) illustrates the use of quantile Fourier regressions for wind power generated at this scale. It is clear that there are seasonal patterns in wind power, with a significant drop in wind generation in the summer. To calibrate the contribution of offshore wind, we were guided by the ISO-New England target for wind generation for 2030. Based on [6, 12], we assume 21687 MW of offshore windpower capacity, corresponding to 1446 IEA-15 MW turbines in the region.

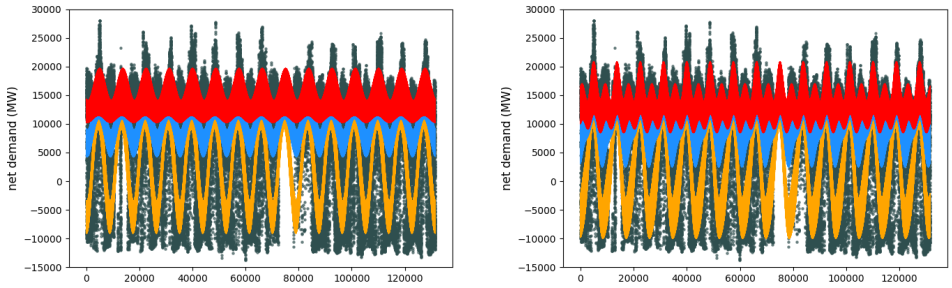

Fig. 4 Quantile Fourier regression with both annual and daily periodicity, fitted to 15 years of net demand data. Left panel: $r = 1$. Right panel: $r = 2$.

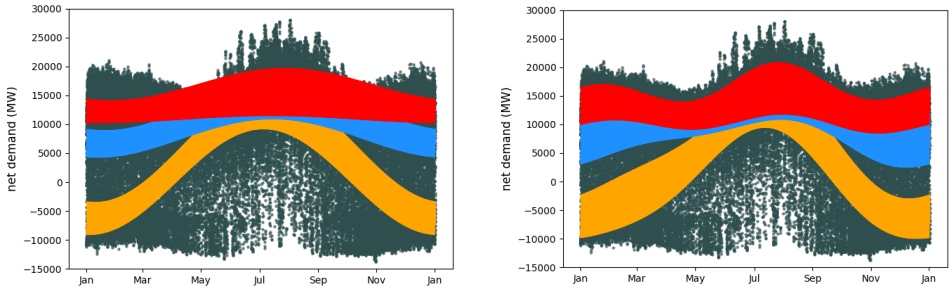

Fig. 5 Quantile Fourier regression with both annual and daily periodicity, fitted to 15 years of net demand data (phase-folded plots). Left panel: $r = 1$. Right panel: $r = 2$.

Since our model is concerned only with the difference between power demand and wind power supply, we combine the two into a single univariate time series of hourly net demand before proceeding to the modelling stage. Figures 4 and 5 juxtapose quantile Fourier regression fits to the net demand series. The improvement obtained by including the second harmonic ($r = 2$) is apparent.

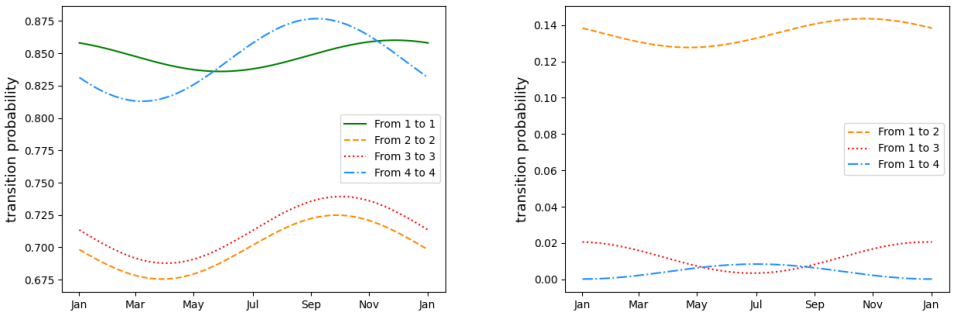

Fig. 6 Fitted transition probabilities for the net demand process. Left panel: transitions from each state to itself. Right panel: transitions from the first (lowest) net-demand state to other states.

The quantile models were then used to model the serial dependence structure as a four-state Markov chain as described in Section 3. All sixteen transition probabilities were permitted to vary annually as simple sinusoids, by using the Fourier basis (2) with $r = 1$. The resulting transition probabilities are illustrated in Figure 6.

Building on the net demand model, we constructed a Markov decision process containing a stylized representation of fourteen combined-cycle gas turbine (CCGT) powerplants with a total of 28000 MW generation capacity. For this stylized example, in each state, we allow increased generation of 2000 MW as "ramp up" action (provided thermal generation level in the current state is 26000 MW or less), decrease of generation by 2000 MW as "ramp down" action (provided current thermal generation is at least 2000 MW), or staying at the same thermal generation level. Equivalently, we have an enormous thermal power plant that can ramp up or down by 2000 MW in each period. The model could be made closer to reality by allowing multiple CCGTs to ramp in each state; this would increase the number of allowable actions in each period significantly.

This example allows us to illustrate the effect of time inhomogeneity in a clear and concise manner. Figure 8 illustrates the optimal operation plans that result from a Markov decision process model in which the decisions may vary by time of day, but not by time of year. In contrast Figure 7 demonstrates optimal decisions for the full model, in which the operation plan may be different on each day of the year. In these figures, red indicates "ramp up" is the recommended action for the state, yellow indicates no change, and green is indicative of "ramp down". Note the significant differences between winter (left panel) and summer (right panel) operations. In particular, the thermal back up operation level rises to cover the afternoon peak demand during the summer, which could be attributed to air conditioning use in hot July afternoons, together with the decrease in wind power generation in the summer. Both operational plans laid out in Figure 7 differ from the all-year-round plan in Figure 8.

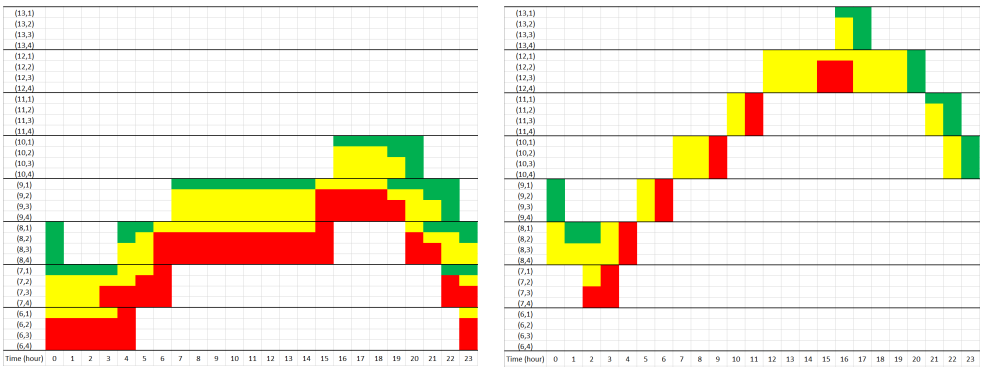

Fig. 7 The offshore wind integration MDP solution. Left panel: optimal decisions on January 1. Right panel: optimal decisions of July 1.

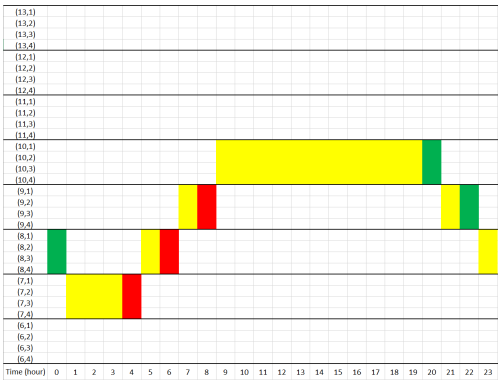

Fig. 8 The offshore wind integration MDP solution: optimal decisions when the decision rule is required to be the same on every day of the year.

7 Acknowledgements

The authors would like to acknowledge the National Science Foundation (GCR award 2020888), The Sloan Foundation (award number 2023-19608) and ISO New England for their generous support of our research.

References

[1] S. Ahmed, A. J. King, and G. Parija, A Multi-Stage Stochastic Integer Programming Approach for Capacity Expansion under Uncertainty. *Journal of Global Optimization* 26 (2003) 3–24.

[2] A. Z. Averbuch, P. Neittaanmäki, and V. A. Zheludev, *Spline and spline wavelet methods with applications to signal and image processing. Volume I, Periodic splines* (2014), Springer.

[3] Betz, A. *Introduction to the Theory of Flow Machines.* (1966) (DG Randall, Trans.) Oxford.

[4] T. Ding, Y. Hu and Z. Bie, Multi-Stage Stochastic Programming With Nonanticipativity Constraints for Expansion of Combined Power and Natural Gas Systems. *IEEE Transactions on Power Systems* 33 no. 1 (2018) 317-328.
[5] T. Ding, M. Qu, C. Huang, Z. Wang, P. Du and M. Shahidehpour, Multi-Period Active Distribution Network Planning Using Multi-Stage Stochastic Programming and Nested Decomposition by SDDIP. *IEEE Transactions on Power Systems* 36 no. 3 (2021) 2281-2292.
[6] EBC. *Presentations: 8th Annual New England Offshore Wind Conference* (2020, December 8). Environmental Business Council of New England. https://ebcne.org/wp-content/uploads/2020/12/Presentations-8th-Annual-New-England-Offshore-Wind-Conference.pdf.
[7] Electricity Authority of New Zealand, *Hydrological Modelling Dataset* (2023).
[8] Federal Energy Regulatory Commission (n.d.). Form No. 714: Annual Electric Reliability Status Report [Data set]. Retrieved May 2023 from https://www.ferc.gov/industries-data/electric/general-information/electric-industry-forms/form-no-714-annual-electric/data.
[9] E. Gaertner *et al*, *Definition of the IEA 15-Megawatt Offshore Reference Wind Turbine* (2020). National Renewable Energy Laboratory. NREL/TP-5000-75698. https://www.nrel.gov/docs/fy20osti/75698.pdf.
[10] W. A. Gardner, A. Napolitano, and L. Paura, Cyclostationarity: Half a century of research. *Signal processing* 86 (2006) 639–697.
[11] F. Hillier, G. Lieberman, *Introduction to Operations Research* 11th edition (2021), MacGraw Hill.
[12] E. Johnson, *New England power system outlook* (2022, June 2). ISO New England. https://www.iso-ne.com/static-assets/documents/2022/06/iso_ne_overview_and_regional_update_cbia_6_2_2022.pdf.
[13] Kirchhoff, Robert H., and F. C. Kaminsky. *Empirical modeling of wind-speed profiles in complex terrain.* No. DOE/ET/10374-82/1. Massachusetts Univ., Amherst (USA). School of Engineering, 1983.
[14] A. J. Kleywegt, A. Shapiro, and T. Homem-de-Mello, The Sample Average Approximation Method for Stochastic Discrete Optimization. *SIAM Journal on Optimization* 12, no. 2 (2002) 479-502.
[15] R. Koenker, *Quantile Regression* (2005), Cambridge University Press.
[16] Massachusetts: *Governor Baker Signs Comprehensive Climate Change Legislation.* (2021, January 2) Massachusetts Office of Energy and Environmental Affairs.
[17] NOAA National Data Buoy Center (n.d.). Station 44025 [Data set]. Retrieved May 2023 from https://www.ndbc.noaa.gov/station_history.php?station=44025.
[18] G. Pritchard, Stochastic inflow modelling for hydropower scheduling problems. *European Journal of Operational Research* 246 (2015), 496–504.
[19] D. Rayner, C. Achberger, and D. Chen, A multi-state weather generator for daily precipitation for the Torne River basin, northern Sweden/western Finland. *Advances in Climate Change Research* 7 (2016) 70–81.
[20] E. Sabet, B. Yazdani, R. Kian, K. Galanakis, A strategic and global manufacturing capacity management optimisation model: A Scenario-based multi-stage

stochastic programming approach. *Omega* 93 (2020).
[21] R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices. *Annals of Mathematical Statistics* 35 (1964) 876–879.
[22] R. Sinkhorn and P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics* 21 (1967) 343–348.
[23] D. Wang and B. J. Adams, Optimization of reservoir operations with Markov decision processes. *Water Resources Research* 22 (1986) 345–352.
[24] D. J. White, *Markov Decision Processes* 1st edition (1993), Wiley.
[25] G. Zakeri, A. B. Philpott, and D. M. Ryan, Inexact Cuts in Benders Decomposition. *SIAM Journal on Optimization* 10, no. 3 (2000) 643-657.