Approximate Dynamic Programming of Long Term Power Generation Schedule

Zhencheng Liang\textsuperscript{1,2*}, Yiming Li\textsuperscript{2}, Yixin Zhuo\textsuperscript{2} and YinSheng Su\textsuperscript{1}

\textsuperscript{1} China Southern Power Grid Dispatching Control Center, Guangzhou, Guangdong, 510623, China
\textsuperscript{2} Guangxi Power Grid Dispatching Control Center, Nanning, Guangxi, 510623, China
*Corresponding author’s e-mail: liangzc@csg.cn

Abstract. With respect to the features in optimal scheduling of long-term power generation such as long cycle, large scale and strong randomness, the paper proposed a multi-stage optimized decision model based on approximate dynamic programming. In this model, the influence of random variables including load, fuel price, and water was tested by means of stochastic simulations. Staged solution was used to reduce the size and difficulty of solving the problem. The proposed value function approximation strategy of pondage and coal inventory solved the "curse of dimensionality" in staged recursion and maintained the overall optimization characteristics after staged decomposition. Through the iterative update between the decision and the approximation function, the approximate optimal decision sequence was solved to obtain the optimization scheme of power generation scheduling. The calculation and analysis results in a real system in some province showed that the modeling by approximate dynamic programming is an effective way to solve such large-scale optimization and has a promising future by virtue of its conciseness, effectiveness, convenience in dealing with random factors, high calculation accuracy and rapidity of solution.

1. Introduction

The optimal scheduling of long-term power generation is within the scope of resource optimization (RO) and involves the optimal allocation of generation resources like hydro and thermal power in a broad range for a long time. It is significant for the industry to achieve energy balance, reduce fossil fuel consumption and improve the utilization rate of clean energy. In electricity market, the department of power dispatching needs to take into account multiple factors such as load demand, fuel price, water and system feature to prepare the long-term power generation plan of hydro and thermal power, the fuel supply plan and the scheduling plan of reservoir.

Based on the deterministic assumption, The literature \[1\] resolved the power into hydro and thermal power by means of system decomposition and coordination to reduce the solution size, and used the coordinating and binding effect of Lagrange multiplier for the iterative solution of major problem's optimal value to achieve the goal of integrative optimization. \[2\] proposed that two-stage dynamic programming should be applied to optimize the pondage, which includes the preliminary idea of value function approximation. \[3\] compared the deterministic model and stochastic model of hydro and thermal power system, and its dynamic programming (DP) and stochastic dynamic programming (SDP) are applicable for problems in small scale. \[4\] adopted Benders' decomposition, so that only small-scale linear programming problems need to be solved at each stage to achieve the goal of solving large-scale scheduling problems with long cycle. \[5\] applied Monte Carlo Simulation to...
analyze the influence of such factors as random inflow, random load and reliability of machine set. [6]considered the coordinated optimization between the generation schedule and the maintenance plan, and its multi-scenario scheduling mode is an effective way to think about the stochastic scheduling problems. As a result, the model can solve a problem more rapidly. It can be seen from the literature review above that the decomposition of problems and the simplification of model are two major ways to study the RO problems. Staged decomposition can divide the problems in the time dimension and is able to take random factors into consideration and reduce the scale of solving a problem before staged decision-making, but it is difficult to ensure the overall optimization characteristics after decomposition. Model linearization considers random factors as a whole and applies simultaneous solution, but the accuracy and efficiency of solution depends on the linearization method and the performance of commercial solver, which restrains its application to some extent. With respect to the large scale programming optimization like RO, the research focus is to consider random factors and solve the problem in a rapid and effective way.

Bellman's Principle of Optimality guarantees the optimization characteristics after staged decomposition, but it is easy to produce "curse of dimensionality" during accurate DP recursion. Professor Powell from Princeton University summarized the research results in stochastic programming and DP, and proposed the modeling framework and solution ideas of Approximate Dynamic Programming (ADP)[7], so as to deal with the programming problems involving determinacy & indeterminacy and continuity & dispersion. ADP To solve the problems in stochastic dynamic programming effectively, it is necessary to realize random variables before staged decision-making based on multi-stage decision model and use approximate structure to approach the function of accurately calculated value in DP. The effective solution of many problems such as dynamic resource optimization[8], large-scale locomotive dispatching [9] and energy planning[10] indicates the application prospect of ADP.

On basis of analyzing traditional RO model, the paper applied the theories of ADP, combined the actual scheduling methods, and established a multi-stage optimization decision model, which set load, fuel price, and water as random variables; set hydro and thermal power, fuel and water consumption as decision variables; and set minimum cost of fuel as the objective. It proposed the value function approximation strategy of pondage and fuel inventory to prevent the "curse of dimensionality" in staged recursion. Following the simulation sampling of value function approximation and random variables, the value function for approximation was iterated and updated in accordance with ADP's solution process to obtain approximate optimal decision solution. At last, the system of hydro and thermal power in a province in South China was modeled and solved to verify the effectiveness of the proposed model and algorithm.

2. Mathematical model of RO

2.1 Objective function
From the perspective of power generation plan and fuel supply plan made by the department of power dispatching, the minimization of fuel cost of all thermal plants was set as the objective function.

\[
\min \sum_{i \in I} \sum_{t \in T} \rho'_t \ g'_i
\]

(1)

where \( T \) is the plan periods; \( I \) is the set of thermal power plant; \( \rho'_t \) is the price of fuel during the period \( t \); \( g'_i \) is the fuel purchased by plant \( i \) during the period \( t \).

2.2 Constraint conditions
Power balance constraint:

\[
\left( \sum_{i \in I} p'_i + \sum_{j \in J} p'_j \right) \Delta t = E'_d
\]

(2)
where $p_{i}^{t}$, $p_{j}^{t}$ represent the power output of thermal plant $i$, hydroplant $j$ respectively; $J$ is the set of hydroplants; $\Delta t$ is the number of hours within the period $t$; $E_d$ is load demand.

The energy consumption characteristic of thermal plant is as follows:

$$F_i^{t} = c_i p_i^{t}\Delta t \times 10^{-6}$$  \hspace{1cm} (3)

$$0 \leq p_i^{t} \leq \bar{p}_i$$  \hspace{1cm} (4)

where $F_i^{t}$ is the fuel consumption of the thermal plant $i$ during the period $t$; $c_i$ is the fuel consumption of thermal plant; $\bar{p}_i$ is the maximum power output of thermal plant $i$.

The dynamic balance of fuel inventory is as follows:

$$G_{i}^{t+1} = G_{i}^{t} + p_{i}^{t} - F_i^{t}$$  \hspace{1cm} (5)

$$G_{i}^{t} \leq G_{i}^{t} \leq \bar{G}_i$$  \hspace{1cm} (6)

where $G_{i}^{t+1}$ is the inventory of the thermal plant $i$ at the end of the period $t$ (the beginning of $t+1$); $G_{i}^{t}$ and $\bar{G}_i$ represent the minimum and maximum inventory of thermal plant respectively.

The generation characteristic of hydroplant is as follows:

$$p_{j}^{t} = b_{h,j}(R_j^{t})^2 + b_{2,j}(Q_j^{t})^2 + b_{3,j}R_j^{t}Q_j^{t} + b_{4,j}R_j^{t} + b_{5,j}Q_j^{t} + b_{6,j}$$  \hspace{1cm} (7)

$$p_j \leq p_{j}^{t} \leq \bar{p}_j$$  \hspace{1cm} (8)

$$Q_j \leq Q_{j}^{t} \leq \bar{Q}_j$$  \hspace{1cm} (9)

where $R_j^{t}$ is pondage of the hydroplant $j$ during the period $t$; $Q_j^{t}$ is the inflow for generation; $b_{h,j}$, $b_{2,j}$, $b_{3,j}$, $b_{4,j}$, $b_{5,j}$ and $b_{6,j}$ are the coefficients of the output characteristic curve of the hydroplant. $\bar{p}_j$ & $\bar{Q}_j$ and $\bar{Q}_j$ are respectively the upper and lower limits of the output and the flow of hydroplant.

The balance of pondage is as follows:

$$R_j^{t+1} = R_j^{t} - 3600\Delta t \left(Q_j^{t} + \sum_{j \in S^u} Q_j^{t} - w_j^{t}\right)$$  \hspace{1cm} (10)

$$R_j \leq R_j^{t} \leq \bar{R}_j$$  \hspace{1cm} (11)

where $Q_j^{t}$ is the outflow of upstream reservoir; $S^u$ is the set of upstream reservoir; $w_j^{t}$ is the water inflow; $\bar{R}_j$ and $\bar{R}_j$ are the upper and lower bounds of reservoir's capacity.

In the hypothesis of traditional model, load, fuel price, and water are certain and known, so it is regarded as a typical nonlinear programming. In the multi-scenario optimization model[6], random variables are realized in the form of scenario and non-deterministic problems are transformed into deterministic problems; the scenario-related deterministic model is then solved to find its statistical regularity.

3. ADP-based solution

3.1 ADP-based RO model

ADP uses the idea of approximate substitution to solve the "curse of dimensionality" related to state variable, decision variable and random variable in DP and SDP. Its main viewpoints include:

1) DP's concept and symbols are used to describe the programming problems, and Bellman's Principle of Optimality is followed for optimization in staged decision-making.
2) Random factors are considered before staged decision-making, and the influence of random variables on decision-making is realized by stochastic simulated sampling.
3) The optimization strategy is open to the methods in other fields, such as random search, simulation and optimization, reinforcement learning, which can be used to select decision strategy. In this way, ADP's range of application is expanded.

ADP provides a powerful framework for solving DP and SDP problems. After the modeling by DP's method, a better decision sequence is obtained by iteration and update.

First, DP's concept and symbol are used to describe the RO problems:

1) State variable 

\[ S_t = (E_t, R_t, G_t, \rho_t) \]

refers to the system state during the period \( t \), including system load, pondage, fuel inventory and price.

2) Decision variable

\[ \chi_t = (p_t, g_t, p_{j_t}, Q_{j_t}) \]

refers to the decision during the period \( t \), including output of thermal plants, volume of purchased fuel, output of hydroplants, generation flow.

3) Random variable

\[ \hat{W}_{t+1} = (\hat{L}_{a+1}, \hat{\rho}_{t+1}, \hat{\hat{W}}_{t+1}) \]

refers to the realization of random variable during the period of time \( t+1 \), including change of load, fluctuation of fuel price, and water from basin.

4) State transition

\[ S_{t+1} = U_{t} (S_t, \chi_t, \hat{W}_{t+1}) \]

refers to the system state during the period of time \( t+1 \); \( U \) represents the system's evolution characteristics. The state transition of RO problems includes the dynamic balance of fuel inventory (5), the dynamic balance of pondage(10), and the realization of random variables.

5) Objective function.

\[
\min_{\pi} \mathbb{E} \sum_{t=1}^{T} C'(S_t, \chi_t(S_t))
\]

where \( C'(S_t, \chi_t(S_t)) \) is the cost of fuel during the period \( t \); the objective function describes the procedure in which the strategy \( \pi \) is selected for decision-making to minimize the expectation on the cost of fuel during the whole periods. \( \chi_t(S_t) \) is the decision function, and \( \Pi \) refers to optional set of strategies.

Meanwhile, the state variable, decision variable, random variable and state transition of the ADP model should satisfy their respective constraints.

3.2 Value function approximation strategy

Based on the features of energy storage of system resources (water and fuel resources), the value function approximation strategy is applied to solve the ADP-based RO model.

\[ V(S_t) \]

is defined as the value function of the system state \( S_t \) to represent the earnings of the resources stored in the system currently. Therefore, let \( \sum_{i \in I} \rho_i g_i \) be the earnings during the period of time \( t \), and transform the solution of minimum cost into the solution of maximum earnings. Suppose that the strategy \( \pi \) is adopted, then the value function of the state during the period of time \( t \) is as follows:

\[
V(S_t) = C(S_t, \chi_t(S_t)) + \sum_{j \in \Pi} C(S_{t+1}, \chi_{t+1}(S_{t+1}))
\]

Based on Bellman's Principle of Optimality, it is further transformed to the recursive solution of multi-stage optimization decision:

\[
V(S_t) = \max \left( C(S_t, \chi_t(S_t)) + \mathbb{E} V(S_{t+1}) \right)
\]

The multi-stage decision process, in which the decision \( \chi_t \) is calculated to maximize the earnings
from the present to the end of period. $E V(S^{t+1})$ is the value function of succeeding state, including random variable $W^{t+1}$, which is represented as the form of expectation. $S_{t}^{'}$ is the state after decision-making, i.e. the state before the realization of random variable at the stage $t+1$ following the decision-making at the stage $t$. At this time, the state and the random variable are separated.

SDP uses the method of Monte Carlo Simulation[9] to calculate the expectation $E V(S^{t+1})$, which is prone to causing "curse of dimensionality". The value function approximation strategy realizes the solution of the decision $x^{t}$ by constructing the approximation function $\tilde{V}(S_{t}^{'})$ of the state after decision-making $S_{t}^{'}$ to replace $E V(S^{t+1})$. The way that the value of approximation function is iterated and updated avoids the difficulty from the accurate calculation of original value function.

With respect to the RO problems including water resource and fuel resource, can be adjusted by pondage and fuel inventory. According to the definition of $S_{t}^{'}$, $\tilde{V}(S_{t}^{'})$ consists of the value function of the pondage of all reservoirs $R_{t}^{i}$ and the fuel inventory $G_{t}^{i}$, which are mutually independent and represented as:

$$\tilde{V}(S_{t}^{'}) = \sum_{j=1}^{J} \tilde{V}(R_{t,j}^{i}) + \sum_{i=1}^{I} \tilde{V}(G_{t,i}^{i})$$

(15)

With regard to the profit value function of resources, $\tilde{V}(S_{t}^{'})$ is the concave function of $R_{t}^{i}$ and $G_{t}^{i}$. Piecewise linear approximation can be used to construct the approximation function[9],[10].

3.3 Solution process

Figure 1. Solution process of the ADP-based RO model

Following the value function approximation and the realization of random variables by stochastic simulated sampling, the solution process of the ADP-based RO model is as shown in figure 1, including:

1) System initialization. System parameter was obtained, the initial value of the value function was set as 0, and the maximum iterations were set as $N$. 

---

**Figure 1. Solution process of the ADP-based RO model**

Following the value function approximation and the realization of random variables by stochastic simulated sampling, the solution process of the ADP-based RO model is as shown in figure 1, including:

1) System initialization. System parameter was obtained, the initial value of the value function was set as 0, and the maximum iterations were set as $N$. 

---
2) Stochastic simulated sampling. The method in [12] was used to achieve the stochastic simulated sampling of the random variables after the $n^{th}$ iteration.

3) Update the state variable. The state variable was determined according to the state transition function, so as to be well prepared for making decision.

4) Solve the staged optimal decision. Based on the system state and the estimation solution equation (15) of value function, interior point method was used for solution in accordance with the non-linear features of some constraints.

5) Update the approximation function. The update samples of value function after making decision[10].

6) Solve the decisions at all stages. An iterative method was used to solve the approximate optimal decision sequence during stochastic process based on the realization of the $N$ stochastic simulated sampling.

7) Outcome. The optimized scheduling results of the power generation plan of hydro and thermal power, the fuel supply plan and the operating plan of reservoir were calculated and obtained according to the approximate optimal decision sequence.

4. Results and discussion

4.1 System description

In the system, a total of 16 thermal plants and 6 hydroplants were included. Given that the prediction deviation of hydro was great from June to December when water inflow was relatively abundant, the standard deviation of water inflow was set as 6.6% on average; given that the prediction deviation of hydroplant was small from January to May when water inflow was relatively sparse, the standard deviation was set as 3.3% on average. The given predicated power generation was set for other hydroplants.

4.2 Result analysis

Based on the method in [13], the traditional model was expanded to multi-scenario optimization model, marked as E_M. The method in [12] was used to generate 500 simulation scenarios with a period of 12 months which included load, fuel price, and water inflow. These simulation scenarios were set as the realization of random variables. The effect of the two methods in dealing with random problems was compared. Fuel cost, fuel consumption and output expectations are shown in the Table 1.

| Model | Cost/10^7Yuan | Fuel/10^7kg | Power/10^8kWh | Time/mins |
|-------|--------------|-------------|---------------|-----------|
| E_M   | 1 608.07     | 1 988.09    | 297.09        | 664.46    | 135       |
| ADP   | 1 612.61     | 1 992.62    | 294.46        | 665.34    | 17        |

The result of E_M was set as baseline. The expectation deviation of fuel cost and consumption were respectively 0.28%, 0.23% through ADP calculation, showing higher calculation accuracy. Its calculation time was only 17 minutes, while the E_M took about 135 minutes. It indicates that ADP not only guarantees higher calculation accuracy, but also greatly improves the calculation speed.

Figure 2. shows the generation plan based on the calculation with the ADP method. The planned output of hydro and thermal power accounts for 37% and 63% respectively. The output of hydropower was determined by water inflow, and the optimization of reservoir was to increase hydropower output at the period of time with large load and high fuel price; the output of thermal power basically synchronized with load, which was consistent with the situation where thermal power was the major power source, and its fuel supply plan could be scheduled on this basis.
Figure 2. Scheduling plan of power generation

The fuel supply plan is shown in table 2. The plan integrates multiple features including generation schedule, fuel price and coal inventory. It is suggested to increase the coal inventory properly when the fuel price is low, and reduce the fuel purchase when the fuel price is high.

|                | 1st quarter | 2nd quarter | 3rd quarter | 4th quarter |
|----------------|-------------|-------------|-------------|-------------|
| Consumption/10^7kg | 558.25      | 396.87      | 456.89      | 416.92      |
| Purchased/10^7kg | 638.85      | 391.07      | 454.84      | 334.14      |
| Difference      | -80.60      | 5.80        | 2.02        | 72.78       |

Optimal operation of hydropower is the key to realize the joint optimization plan of hydro and thermal power. The water level scheduling curve of hydroplant Y is shown in figure 3. During the dry season (from January to June), the water inflow and the water stored last year were used for power generation. During the flood season (from July to October), the impoundment ability of reservoir was used to adjust the water level according to water inflow. The upstream water level should be lifted to increase the water head and reduce the blocking for generation. At this time, the power plant had a strong ability of peak clipping by using capacity of reservoir.

Compared with hydroplant Y, the regulating capacity of reservoir in hydroplant J is larger. The plant has strong impoundment ability during the flood season, and can reserve the water for future use during the dry season. The figure 3 shows from January to April when the output of hydroplant Y was small during the dry season, hydropower J used the stored water for power generation and bore more load from the power grid.

Figure 3. Scheduling water level of hydropower Y and Y hydropower J

It can be seen from the average flow in the basin for many years that the flood season of other basins was earlier than that of the Red River Basin, so it is possible to carry out compensation scheduling across basins. The figure 4 shows that from February to May when the output of the Red River Basin was restricted by water inflow, the output of other basins could be used for partial compensation.
5. Conclusion

The paper applied the ADP method to model and solve the RO problems including hydro and thermal power, and utilized ADP's staged optimization features and ability in applying stochastic simulated sampling to deal with random variables. By theoretical analysis and case study, the following conclusions are made:

1) The value function approximation strategy applies to the multi-stage decision model of RO problems. By using the approximation function of pondage and fuel inventory, the strategy achieves the decoupling between stages, reduces the size of problems, improves the solution efficiency and maintains optimization characteristic.

2) The method of ADP modeling can effectively deal with random factors. Compared with the methods of multi-scenario model, ADP realizes the random variables by using stochastic simulated sampling, applies samples to update value function, and obtains the approximate effect when random factors are considered.

3) The proposed ADP-based RO model gives an overall consideration to multiple factors such as load, fuel price, water and system feature; prepares the long term water utilization and power generation plan for hydropower; achieves the plans including the water optimization plan of reservoir, the cascade optimization plan and the coordinated complementation plan between basins; and prepares the fuel purchase plan to minimize the cost based on the power generation plan of thermal plant and the prediction plan of fuel price.

References

[1] Han Xueshan, Liu Zhuo. Decoupled economic dispatch including unit ramp rate and network security constraints[J]. Automation of Electric Power Systems, 2002, 26(13): 32-37(in Chinese).

[2] Wu L, Shahidehpour M, Li T. Stochastic security-constrained unit commitment[J]. IEEE Transactions on Power Systems, 2007, 22(2): 800-811.

[3] Ferrero R W, Rivera J F and Shahidehpour S M. A dynamic programming two-stage algorithm for long-term hydrothermal scheduling of multireservoir systems[J]. IEEE Transactions on Power Systems, 1998, 13(4): 1534-1540.

[4] Zambelli M, Siqueira T G and Cicogna M, et al. Deterministic versus stochastic models for long term hydrothermal scheduling[C]. Power Engineering Society General Meeting, 2006. IEEE.

[5] Pinto R J, Borges C L T and Maceira M E P. An efficient parallel algorithm for large scale hydrothermal system operation planning[J]. IEEE Transactions on Power Systems, 2013, Page(s): 1-9.
[6] Baslis C G, Papadakis S E, Bakirtzis A G. Simulation of optimal medium-term hydro-thermal system operation by grid computing[J]. IEEE Transactions on Power Systems, 2009, 24(3): 1208-1217.

[7] Powell W B. Approximate dynamic programming : Solving the curses of dimensionality[M]. A JOHN WILEY&SONS, INC., New York. 2011.

[8] Topaloglu H, Powell W B. Dynamic programming approximations for stochastic time-staged integer multicommodity-flow problems[J]. INFORMS Journal on Computing, 2006, 18(1): 31-42.

[9] Powell W B, Bouzaiene Ayari B, Cheng C, et al. Strategic, tactical and real-time planning of locomotives at Norfolk Southern using approximate dynamic programming[C]. Joint Rail Conference. American Society of Mechanical Engineers, 2012: 491-500

[10] Powell W B, George A, Simão H, et al. Smart : A stochastic multistate model for the analysis of energy resources, technology and policy[J]. INFORMS Journal on Computing, 2012, 24(4): 665-682.

[11] Pereira M V F, Pinto L M V G. Multi-stage stochastic optimization applied to energy planning[J]. Mathematical Programming, 1991, 52(1-3): 359-375.

[12] Dupacova J, Consigli G, Wallace S W. Scenarios for multistage stochastic programs[J]. Annals of Operations Research, 2000(100): 25-53.

[13] Antonio J.Conejo, Miguel Carrion, Juan M. Morales. Decision making under uncertainty in electricity markets[M]. Springer, 2010.