

# 强化学习方法在水火混杂 AGC 系统中的应用

李红梅, 严 正

(上海交通大学电子信息与电气工程学院, 上海市 200240)

**摘要:** 多区域互联系统通常是水火电自动发电控制(AGC)机组并存系统,系统的非线性因素及复杂性使得基于线性理论的比例积分调节方式不能适应智能电网的要求。文中针对水轮机系统的非线性特点,将环境知识转化成强化学习算法的先验知识以加快 AGC 的调整速度。在此基础上,完整阐述了水火电混杂的多区域系统基于强化学习算法的自动发电智能化控制设计思想。为使模型更具一般性,建立了三区域混杂系统仿真模型,仿真结果验证了该方法的可行性。

**关键词:** 自动发电控制(AGC); 水火混杂系统; 非线性; 先验知识; Q 学习

## 0 引言

传统的负荷频率控制,即狭义的自动发电控制(AGC)采用的控制信号是将区域控制偏差(ACE)进行积分后送到二次调频回路调整发电机组出力。基于线性理论进行分析时,由于控制回路中采用负反馈,因此,不论积分增益系数如何设置,比例积分环节都能够将误差消除从而达到最终的稳定状态<sup>[1]</sup>。这是传统的 AGC 理论的基础<sup>[2-3]</sup>。但在实际系统中,用线性控制原理去指导实际 AGC 系统,其动态性能无法令人满意<sup>[4-5]</sup>。其原因就是由于无论是水电系统还是火电系统都存在非线性环节,且二次控制回路的比例积分增益系数  $k_i$  不仅与系统结构参数有关,还与负荷变化相关<sup>[1]</sup>,因此,  $k_i$  的整定应随着系统负荷的变化实时调节,才能满足负荷变化时频率的变化要求。水火电机组都面临着如何设定积分增益系数以满足系统实时变化需求的问题。

文献[6]中,用具有先验知识的 Q 学习方法设计了火电机组自动发电控制器:将非线性的 AGC 系统离散化成 Markov 链;ACE 作为状态变量;用 Q 学习方法控制器代替传统的比例积分控制器,为了进一步加快学习速度,用模糊综合决策的方法获得火电机组 AGC 环境信息,为强化学习提供先验知识。

但是,在多区域系统中,不仅包括火电机组还包括众多的水电机组。为了全面考虑,本文侧重描述水电机组 AGC 控制器设计,突出水电机组 AGC 的特点及如何获得水电机组 AGC 的环境知识,并结合文献[6]中火电机组控制器的设计提出了水火混

杂系统 AGC 控制器设计思想。最后采用典型的三区域系统模型进行仿真验证。由于文献[6]已经对火电机组 AGC 控制器进行了详细描述,本文侧重描述的是水电机组 AGC 控制器设计,结合文献[6]给出了完整混杂系统的仿真结果。

## 1 水轮机 AGC 系统中的非线性环节

文献[6]只针对火电机组 AGC 系统中存在的非线性环节进行阐述,主要包括延时、发电机调节容量限制和机组爬坡率限制、死区等。本文针对水轮机系统中影响 AGC 的非线性环节进行介绍。

水轮机调节系统是一个非线性时变系统,主要是因为电液伺服控制系统(或机械液压控制系统)存在非线性及其引水系统自身存在非线性。其引导阀辅助接力器和主配压阀接力器都存在起限幅作用的饱和非线性环节。水轮机调节系统的非线性直接影响发电机输出电能频率和过渡过程中的动、静态特性,是不利于系统稳定的重要因素<sup>[7]</sup>。水流水头因水电站类型不同而存在很大区别,本文假设引水系统为刚性水锤,即不考虑水头因子和水流不稳定作用。

### 1.1 水轮机 AGC 系统的非线性约束

#### 1.1.1 延时

为了实时跟踪负荷变化以调整机组出力,SCADA 系统需要收集相关的系统状态数据并将遥调指令下发到各个发电厂或直接下发到机组。因此,信息的采集分析、命令的下发和执行都需要一定的时间来完成。本文只考虑由上述各个环节所造成的固定延时。

#### 1.1.2 水轮机调节系统非线性环节

1) 频率死区设定。为了在电网频率变化较小情况下提高机组运行的稳定性,调速器一般都有频率

死区设定。当频率偏差小于死区设定值时,调速器停止动作。本文中,水轮机调速器死区的设定值与汽轮机相同,为 $\pm 0.033 \text{ Hz}$ <sup>[8]</sup>。

2) 调速器机械液压系统中的非线性。以往采用的水轮机调速器以机械液压调速器居多,现今采用较广泛的是电液调速器,主要区别是后者更灵活,便于使用<sup>[9-10]</sup>。因为两者功能基本相似,在此只讨论机械液压系统调速器的非线性问题。在机械液压系统中主要考虑接力器关闭时间和关闭特性,关闭时间是接力器从100%开度关闭到0%开度的最短时间,它限制着接力器的关闭速度,其特性类似于火电机组爬坡率约束<sup>[10]</sup>。因此,文献[7]特别强调指出,机械液压系统模块中必须加入饱和度模块和比例限幅模块来表示系统固有的非线性特性。

## 1.2 非线性因素影响比较

为了比较非线性因素对频率变化的影响程度,将上述非线性环节即延时、死区、机械液压系统中的比例限幅模块分别加入传统 AGC 模型中,其他参数不变情况下的仿真结果可参照图 1。仿真过程中,假设积分增益系数  $k_i = 0.3$ , 延时取 2 s, 调速器死区设置为 0.033 Hz, 比例限幅值为每秒 0.01 (标幺值), 负荷扰动相同。

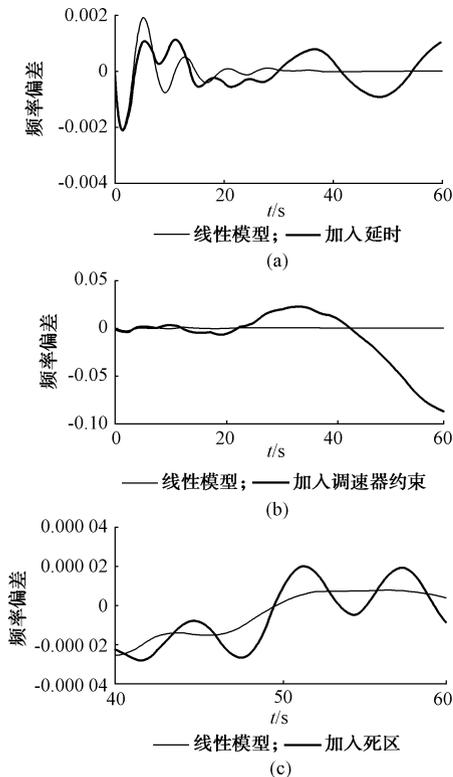


图 1 系统非线性因素对频率的影响  
Fig. 1 Influence of non-linearity factors on frequency variation

从仿真结果可以看出,对频率影响最大的是调速器的比例限幅环节(见图 1(b)),其次是延时的影响(见图 1(a)),然后是死区(见图 1(c))。其中,死区的影响体现在当频率偏差小于调速器死区设定时,会在小范围内引起频率波动,这种波动会增加发电机无谓的出力调节,从而增加机械磨损。

## 2 水火电混杂系统 AGC 控制器强化学习步骤

AGC 系统是一个周期性控制系统,因此可以根据其控制周期进行时间离散化处理,即 AGC 系统在这个时刻所采取的措施只与当前状态有关而无需考虑系统的初始状态以及前一个状态如何,这是一个典型的 Markov 链。基于此,可以用强化学习控制器代替传统二次调频控制中的比例积分环节,将 ACE 作为系统状态量,将机组出力调整值作为行动值。实际系统中对 ACE 的设定有一个范围,即当区域联络线误差大于某一个值时,AGC 机组挂机不动作,交由其他控制系统去处理,例如切负荷;当 ACE 小于某一个数值时,系统认为这是一个无意识交换量,AGC 机组也不必动作。因此,很容易将 ACE 值在有限的状态空间中进行离散化处理。机组出力即行动值也受机组最大出力及最小出力约束,也可以在有限的状态空间中进行离散化处理。不难看出,AGC 控制算法实际上就是一个从 ACE 构成的状态空间到机组出力变化的映射,这个映射就称为策略。

Q 学习是一种没有环境知识的动态决策过程,学习过程缓慢。为了改善 AGC 控制器的初始学习状态,在文献[6]中介绍了基于先验知识的 Q 学习方法,它是借助于模糊综合决策的方法将专家经验和环境知识转化成 AGC 控制器学习的先验知识。

用于评估 AGC 系统非线性的因素集已经在第 1 节加以介绍,并比较了每个非线性因素对系统的影响。调整 AGC 控制策略时如果能够充分考虑这些因素对决策的影响,将更有利于选取正确的行动。因此,模糊综合决策方法被用来评估这些非线性因素对系统的影响<sup>[11]</sup>。其中,决策集就是强化学习行动集。模糊权重集表示每一个因素对决策的影响,权重可以通过经验直接得到。因此,可以通过因素集、决策集和模糊权重集获得综合评估结果,这就是 Q 学习的先验知识。基于先验知识的 Q 学习步骤如下。

步骤 1: 建立状态集、决策集和因素集

1) 状态集  $X = \{x_1, x_2, \dots, x_{11}\}$  由离散化的 ACE 值组成。将 ACE 值离散化成 11 个区间值:  $0.020 \sim 0.016$ ;  $<0.016 \sim 0.012$ ;  $<0.012 \sim 0.008$ ;

$< 0.008 \sim 0.004$ ;  $< 0.004 \sim 0.001$ ;  $< 0.001 \sim -0.001$ ;  $< -0.001 \sim -0.004$ ;  $< -0.004 \sim -0.008$ ;  $< -0.008 \sim -0.012$ ;  $< -0.012 \sim -0.016$ ;  $< -0.016 \sim -0.02$ 。其中,ACE 死区值代表的是最小 ACE 值,即 AGC 不需要动作的最小区域控制误差值,其范围是  $0.001 > x_6 > -0.001$ 。参数用标么值表示。

2) 决策集,即强化学习的行动集,是 AGC 系统中每个控制周期内机组调整的出力值。由于受机组爬坡率约束,每一个周期的最大行动值  $\Delta P_{\max}$  可以设定为周期内机组可能的最大技术出力上限值。汽轮机最大行动值  $\Delta p_t$  设为 0.002 5,水轮机最大行动值  $\Delta p_h$  设为 0.005。将其离散化成 7 个区间,则汽轮机和水轮机的行动集可表示如下:

$$\left\{ \begin{array}{l} \Delta P_t = \left\{ 0.0025, \frac{0.0050}{3}, \frac{0.0025}{3}, 0, \right. \\ \left. \frac{-0.0025}{3}, \frac{-0.0050}{3}, -0.0025 \right\} \\ \Delta P_h = \left\{ 0.005, \frac{0.010}{3}, \frac{0.005}{3}, 0, \frac{-0.005}{3}, \right. \\ \left. \frac{-0.010}{3}, -0.005 \right\} \end{array} \right.$$

3) 汽轮机的因素集由机组最大出力约束(GRC)、爬坡速率约束(URC)、延时和调速器死区构成。GRC 可划分成 3 个区间值,分别为大于出力最大上限值、小于最小出力下限值、介于两者之间<sup>[1]</sup>。调速器死区设置为  $\pm 0.033$  Hz。汽轮机爬坡速率一般设为每分钟 2%~5%,即每个周期(AGC 周期设置为 2 s)的 URC 为 0.001~0.003。水轮机的因素集包括延时、死区和调速器机械液压系统中的饱和度和比例限幅约束。饱和度模块与汽轮机的 GRC 相似,设定为从 1 到 -1 的区间段,机械液压变化速率设置为每 2 s 变化 0.005(标么值)。死区设置与汽轮机相同。

步骤 2: 依据专家经验建立模糊评估矩阵和权重集。依据第 1 节的叙述,可以将因素集进行排序,对系统频率变化影响最大的是 URC,然后是延时,随后是死区。这个次序是确定因素集权重的重要依据。利用模糊综合决策方法计算状态集  $X$  中的所有状态  $x_i$  值,按总体目标综合衡量采取各个决策的优越度向量  $B_i$ 。将各个状态  $x_i$  下的  $B_i$  进行“归一”调整,其结果作为 Q 学习的先验知识,对状态  $x_i$  下的 Q 值进行初始化。

步骤 3: 开始 Q 学习。在周期  $t$  内,根据当前状态  $x_t$ ,选择动作  $a_j$ ,到达新的状态  $x_{t+1}$ ,以及获得一个立即回报  $r(x_t, a_j)$ ,更新 Q 值,重复该过程直到学习结束。

### 3 混杂系统 AGC 仿真结果分析

水火混杂三区域系统 AGC 模型见附录 A 图 A1,其中区域 2 为水电机组。二次调频系统控制器采用 Q 学习算法代替传统的比例积分控制器。在控制系统模型中,非线性因素被充分考虑并表示出来。3 个区域的控制器独立运行,即强化学习智能体之间没有耦合。所有参数采用文献[12]的数据。

仿真过程中采取的措施如下:

1) 用 MATLAB/Simulink 中的白噪声模仿系统负荷扰动,负荷在 0.02 到 -0.02 之间随机扰动。

2) 为便于比较,方法 A 中的 Q 函数初始值随机生成,而在方法 B 中的 Q 函数初始值是基于方法 A,然后将环境先验知识加入 Q 函数的初始值中。方法 A 和 B 中的负荷扰动均发生在区域 3。

3) 机组爬坡速率约束着机组在每个周期的最大调整出力值,为了比较爬坡速率大小对 AGC 系统的影响,方法 A 和 B 中水轮机和汽轮机的最大行动值均设置相等,为  $\Delta P_t^{\max} = \Delta P_h^{\max} = 0.0025$ 。事实上,水轮机可以在几分钟内就达到满出力,因此方法 C 中设置水轮机最大出力为方法 B 的 2 倍,即  $\Delta P_t^{\max} = 0.0025, \Delta P_h^{\max} = 0.005$ 。在方法 C 中,负荷扰动发生在区域 2。

4) 学习结束的判据是 3 个区域的 Q 函数前后 2 次差值小于给定误差值,且频率偏差值在允许范围之内。

表 1 列出了 3 种方法比较结果。从表中数据可以得出以下几个结论:

1) 如果设置 Q 函数前后误差值小于 0.005 结束学习时,方法 A 需要 1 450 幕(episode)的学习,而方法 B 由于有先验知识,因此只需要 840 幕学习。方法 C 的最大行动值是方法 B 的 2 倍,因此只需要 609 幕就能达到要求。

2) Q 函数误差值从 0.005 减少到 0.001 时,方法 A 和 B 的学习幕数基本相同,都经历了大约 700 幕,而方法 C 只需要 400 幕;从 0.001 0 减少到 0.000 2 时基本都需要 200 幕的学习。这是因为方法 C 将机组出力调整到 0.02 只需要 4 个周期就可以完成,而方法 A 和 B 需要 8 步。

表 1 不同仿真方法学习幕数比较  
Table 1 Comparison of learning episodes

学习误差 $\Delta Q$	学习幕数		
	方法 A	方法 B	方法 C
0.005 0	1 450	840	609
0.001 0	2 353	1 510	1 067
0.000 2	2 660	1 714	1 235

方法 B 和 C 的仿真结果如图 2、图 3 所示。由图 2 可见,在区域 3 为了响应负荷 0.02 的扰动,控制器经过 8 个周期来调整机组出力,且每个周期都采用最大出力值 0.002 5。区域 1 和 2 由于不需要对区域 3 发生的扰动做出响应,因此机组出力没有变化。在方法 C 中,由于水轮机最大行动值为 0.005,因此只需要 4 个周期就可以调整机组出力来平衡负荷变化。从图 2(b)和图 3(b)的曲线变化可以看出,由于方法 C 中水轮机调整没有方法 B 中汽轮机平缓,因此引起频率波动也较大,需要达到平缓的时间也比较长。

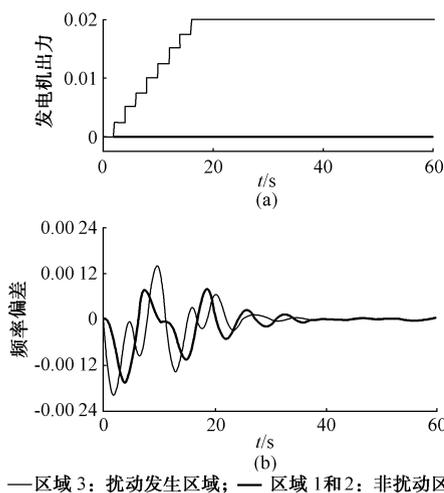


图 2 火电机组区域负荷扰动时系统频率变化及机组出力情况

Fig. 2 Frequency and power output varying with load disturbance happened in area of thermal power system

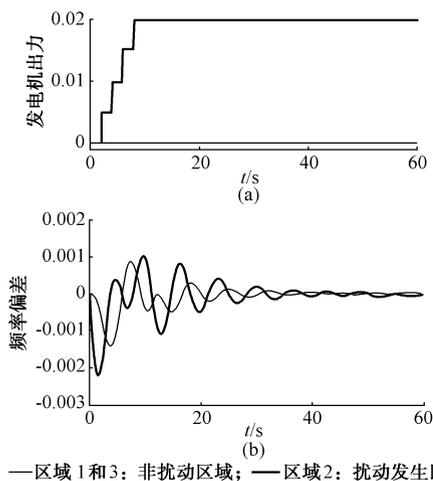


图 3 水电区域负荷扰动时系统频率和机组出力变化情况

Fig. 3 Frequency and power output varying with load disturbance happened in area of hydro power system

## 4 结语

1)学习方法中加入先验知识能有效改善学习过程,而当 Q 学习达到一定程度时,由于智能体积累了一定的经验,因此学习速度也会加速。

2)尽管本文采用了较复杂的水火电混杂系统,但由于每个区域的控制器是独立学习的,相互之间没有耦合,也简化了算法。

与文献[1]相比,本文中采用的模型更复杂,且充分考虑了水电机组和火电机组所特有的非线性环节的特点,因此也更符合实际混杂系统的变化,能够满足实际系统需求。用智能化的学习控制器代替传统的比例积分控制将是 AGC 系统在智能化电网要求下的一个可能的新的发展方向。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>)。

## 参考文献

- [1] 熊观佐. 互联系统分析与运行. 北京:水利电力出版社,1995.
- [2] ELGERD O I, FOSHA C E. Optimum megawatt frequency control of multi-area electric energy systems. IEEE Trans on Power Apparatus and Systems, 1970, 89(4): 556-563.
- [3] NANDA J, KAUL B. Automatic generation control of an interconnected power system. IEE Proceedings-C, 1978, 125(5): 385-390.
- [4] HIYAMA T. Applications of rule based stabilizing controller to electrical power systems. IEE Proceedings: Generation, Transmission and Distribution, 1989, 136(3): 175-181.
- [5] DASH P K, LIEW A C, MISHRA B R. An adaptive PID stabilizer for power systems using fuzzy logic. Electric Power Systems Research, 1998, 44(3): 213-222.
- [6] 李红梅,严正. 具有先验知识的 Q 学习算法在 AGC 中的应用. 电力系统自动化,2008,32(23):36-40.  
LI Hongmei, YAN Zheng. Application of Q-learning approach with prior knowledge to non-linear AGC system. Automation of Electric Power Systems, 2008, 32(23): 36-40.
- [7] 罗旋. 水轮机调节系统的研究与仿真[D]. 北京:华中科技大学, 2005.
- [8] 魏守平,伍永刚,林静怀. 水轮机调速器与电网负荷频率控制. 水电自动化与大坝监测,2006,30(1):18-22.  
WEI Shouping, WU Yonggang, LIN Jinghuai. Hydro-turbine governor and grid load frequency control: Part two simulation of grid load frequency control. Hydropower Automation and Dam Monitoring, 2006, 30(1): 18-22.
- [9] KUNDUR P. Power system stability and control. New York, NY, USA: McGraw Hill, 1994.
- [10] 魏守平,伍永刚,林静怀. 水轮机调速器与电网负荷频率控制: (一)水轮机控制系统的建模及仿真. 水电自动化与大坝监测, 2005,29(6):18-22.  
WEI Shouping, WU Yonggang, LIN Jinghuai. Hydro-turbine governor and grid load frequency control: Part one modeling

and simulation of hydro-turbine governor control system. Hydropower Automation and Dam Monitoring, 2005, 29(6): 18-22.

- [11] 杜春霞,高云,张文. 多智能系统中具有先验知识的 Q 学习算法. 清华大学学报:自然科学版,2005,45(7):981-984.  
DU Chunxia, GAO Yun, ZHANG Wen. Q-learning with prior knowledge in multi-agent systems. Journal of Tsinghua University: Science and Technology, 2005, 45(7): 981-984.
- [12] NANDA J, MANGLA A. Automatic generation control of an interconnected hydro-thermal system using conventional integral and fuzzy logic controller// Proceedings of IEEE

International Conference on Electric Utility Deregulation, Restructuring and Power Technologies, April 5-8, 2004, Hong Kong, China: 372-377.

---

李红梅(1974—),女,通信作者,博士研究生,主要研究方向:电力系统自动控制、电力系统优化运行。E-mail: micelle0126@sjtu.edu.cn

严正(1964—),男,教授,博士生导师,主要研究方向:电力系统优化运行、电力系统稳定分析、电力市场。E-mail: yanz@sjtu.edu.cn

## Application of Reinforcement Learning Method in a Hydro-thermal Hybrid Automatic Generation Control System

*LI Hongmei, YAN Zheng*

(Shanghai Jiaotong University, Shanghai 200240, China)

**Abstract:** Normally, the hydro and thermal automatic generation control (AGC) units exist simultaneously in a multi-area interconnected system. Owing to the inherent non-linearity and complexity of the system, it's hard for an AGC system based on classical linear control theory to meet the requirements for smart grid development. In light of the non-linear characteristics of the hydro unit system, the reinforcement learning method is employed to transform AGC environment knowledge into prior knowledge to develop the AGC controller of a hydro-unit. A detailed analysis is made of the AGC controller design for the multi-area hybrid system of a hydro-power unit and thermal-power unit. In order to generalize the model proposed, a typical three-area hybrid system model is developed with the feasibility of the design method shown by simulation results.

This work is supported by National Natural Science Foundation of China (No. 90612018).

**Key words:** automatic generation control (AGC); hydro-thermal hybrid system; non-linearity; prior knowledge; Q-learning

# 附录A

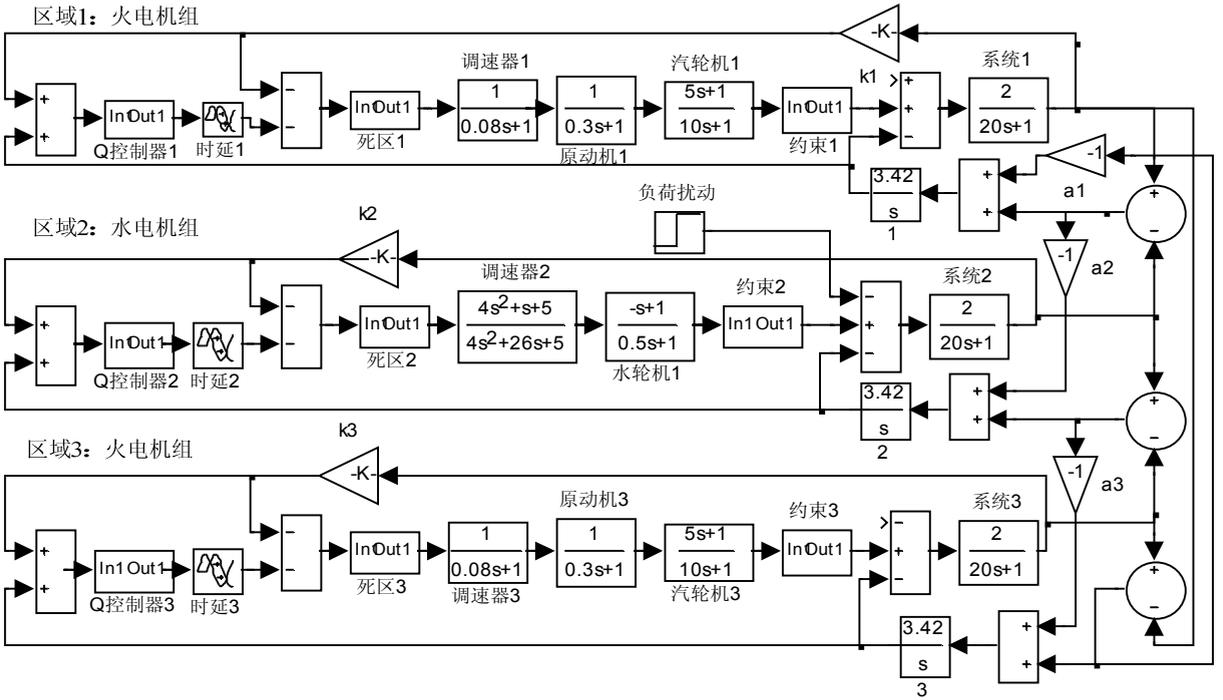


图 A1 三区域混杂系统 AGC 控制模型

Fig. A1 Transformation function model of an interconnected triple-area hydro-thermal system