

Deterministic and Stochastic Optimal Control for Batch Cooling Crystallization

Tushar Gupta

Abstract

Minimization of operation costs and the enhancement in product quality have been major concerns for all industrial processes. The field under study here is batch crystallization which is affected heavily by the uncertainties in measurements and other errors. This work analyses various optimization approaches for the batch crystallization process that are robust to model error. All of the methods involve maximising an objective function by manipulating the cooling profile. At first, the Deterministic approach uses experimental kinetic parameters, which is then extended to Stochastic optimization to incorporate uncertainties in them. Lastly, a novel approach named Polynomial Chaos Expansions is implemented which has been applied successfully to other domains for Nonlinear Model Predictive Control but was not explored in detail in the field of batch cooled crystallization. It successfully includes probability distributions for the parameters into the model to provide a more robust optimization strategy.

Keywords: Stochastic Optimal Control, Polynomial Chaos Expansions, Robust Optimization, Batch Crystallization, Predictive Control, Optimum Temperature Profile

1. Introduction

Numerous industries today, such as pharmaceutical, chemical, photographic etc. employ the batch crystallization process for the preparation of crystalline products with high degree of purity. A common goal of each crystallization process is to obtain a narrower Particle size distribution (PSD) of the desired product. The PSD has a strong influence on the downstream processing and, hence, reproducible PSD in each operation is of prime importance. Thus, finding an effective control strategy to obtain the resulting crystals with a desired Crystal Size Distribution becomes significant in order for improving the performance of both the batch crystallization process and the subsequent processes which depend on it.

Crystallization is the (natural or artificial) process where the atoms or molecules are highly organized into a solid structure known as a crystal. Some of the ways which crystals form are through precipitating from a solution, melt or more rarely deposited directly from a gas. In order for crystallization to take place a solution must be "super-saturated". **Supersaturation**(ΔC) is a condition in which the solute concentration in the solution is higher than the solubility. It acts as the driving force for the crystallization process and affects the final quantity of product formed. It is mathematically expressed as :

$$\Delta C = C - C_s \quad (1)$$

where C_s is the concentration of the solute in the saturated solution. In the following work, the method in focus is cooling crystallization in which superstauration magnitude is determined by the cooling rate. Thus, determination of an optimal cooling rate or a temperature trajectory becomes the objective of the current study.

This work formulates and analyses various control strategies for a cooling crystallization process represented through the population balance equation. **Deterministic Optimal Control** aims at finding the an optimum temperature profile to maximise an objective function selected to achieve a desired volume of the product. Herein, the experimental kinetic parameters are employed to simulate a batch crystallization process. **Stochastic Optimal Control** undertakes the task of quantifying the uncertainities which creep in due to experimentation. It aims to achieve a maximum expected value for the desired product, simultaneously incorporating randomness in the process parameters into the model. Namely, two methods **Ito Process** and a novel approach **Polynomial Chaos Expansions** are employed for this purpose.

Most of the reported works in the this field deal with the determination of optimal temperature or supersaturation trajectory for the batch crystallizer. The concept of programmed cooling in batch crystallizers was first discussed by Mullin and Nyvlt [?] in 1971. Later, in 1974, A. G. Jones [?] presented a mathematical theory based on moment transformations of population balance equations. He used the continuous maximum principle to predict optimal cooling curves. Rawlings et al. [?] discussed issues in crystal size measurement using laser light scattering experiments and optimal control problem formulation. In 1994, Miller and Rawlings [?] discussed the uncertain bounds on model parameter estimates for a batch crystallization system. ost importantly optimal temperature prediction for batch crystallization has also been done by Hu et al.hu, Shi et al.[?], Paengjuntuek et al.[?], and Corriou and Rohani.[?], the data and knowledge from which have been used in further work in this project.

Stochastic modeling of particulate processes and parameter estimation using the experimentally measured particle sizes has attracted many researchers. Grosso et al.[?] presented a stochastic approach for modeling PSD and comparative assessments of different models. Ma et al.[?] presented a worse-case performance analysis of optimal control trajectories by considering features such as the computational effort, parametric uncertainty and control implementation inaccuracies. Monte Carlo simulations have also been used to propogate uncertainties but often present the problem of high computational demand, for which approximations are proposed. Nagy and Braatz (2007) [?] have shown that Polynomial Chaos Expansions(PCE) is a computationally efficient alternative to Monte Carlo simulations for propagating uncertainty in dynamic models. PCE is based on orthogonal basis functions thus requiring smaller function evaluations for the calculation of numerical integrations needed for obtaining statistical moments. The computational advantages of PCEs for robust control and optimization has been shown by Nagy and Braatz [?], Kim et al.[?], Kumar and Budman[?].

The focus of the current research activity is to incorporate parametric uncertainties in the mathematical formulations of batch crystallization process for building a robust model. In the deterministic approach, kinetic parameters from the experimental data have been used to model the system. Next, stochastic Ito processes are used to assimilate the errors in the experimental data. Finally, PCE are demonstrated as an effective and novel method to achieve the desired objective function value. A case study of an unseeded crystallization process is also included to authenticate the methodology.

2. Mathematical Background

This section introduces the underlying concepts needed to build a computational model of a crystallization process which involves regulating the population of particles, also termed as a particulate process. The population is described by the density of a suitable extensive variable, usually the **number of particles**, but sometimes by other variables such as the mass or volume of particles. The usual transport equations expressing conservation laws for material systems apply to the behavior of single particles. Particulate processes are characterized by properties such as particle shape, size, surface area, mass, and product purity.

A population balance formulation describes the process of crystal size distribution with time most effectively. Thus, modeling of a batch crystallizer involves the use of population balances to model the crystal size prediction and the mass balance on the system can be modeled as a simple differential equation having concentration as the state variable. The population balance can be expressed as eq :

$$\frac{\partial n(r, t)}{\partial t} + \frac{\partial G(r, t)n(r, t)}{\partial r} = B \quad (2)$$

where **n** is the number density distribution, **t** is the time, **r** represents the characteristic dimension for size measurements, **G** is the crystal growth rate, and **B** is the nucleation rate. Both growth and nucleation processes describe

crystallization kinetics, and their expression may vary, depending on the system under consideration.

In this work, the system under consideration is potassium sulfate, which has been studied earlier for its kinetics by Hu et al. [?], Shi et al. [?], and Paengjuntuek et al.[?]. The expressions are given as follows:

Nucleation kinetics⁽⁵⁻⁷⁾ are defined by :

$$B(t) = k_b \exp(-E_b/RT) \left(\frac{C - C_s(T)}{C_s(T)} \right)^b \mu_3 \quad (3)$$

Growth Kinetics⁽⁵⁻⁷⁾ are given by:

$$G(t) = k_g \exp(-E_g/RT) \left(\frac{C - C_s(T)}{C_s(T)} \right)^g \quad (4)$$

where k_b and k_g are constants of the system, E_b and E_g are activation energies, and b and g are exponents of nucleation and growth, respectively. The values for these kinetic parameters for the given system have been mentioned in Table1. The following equations are used to evaluate the saturation(C_s) and metastable(C_m) concentrations corresponding to the solution temperature T (expressed in units of °C)[?].

$$C_s(T) = 6.29 \times 10^{-2} + 2.46 \times 10^{-3}T - 7.14 \times 10^{-6}T^2 \quad (5)$$

$$C_m(T) = 7.76 \times 10^{-2} + 2.46 \times 10^{-3}T - 8.1 \times 10^{-6}T^2 \quad (6)$$

The mass balance, in terms of concentration of the solute in the solution, is expressed as :

$$\frac{dC}{dt} = -3\rho k_v G(t) \mu_2(t) \quad (7)$$

where ρ is the density of the crystals, k_v the volumetric shape factor, and μ_2 is the second moment of particle size distribution (PSD). Since $n(r, t)$ represents the population density of the crystals, the i -th moment of the particle size distribution(PSD) is given by :

$$\mu_i = \int_0^\infty r^i n(r, t) dr \quad (8)$$

The above equations along with the Population Balance Equation(PBE) represent a complete model of a seeded batch crystallizer. Population balance equations are multidimensional, which poses a problem with their implementation in complex control functions, hence use of a model order reduction becomes imperative.

For simplification, we reduce the population balance equations into **Moment balance equations** which has been established as an efficient method by Yenkie et al.[?]. This is done by multiplying the equation (2) with r^i on both sides to generate the expression given by equation (8). Converting the model into Ordinary differential equations proves to be advantageous, since it is difficult and time-consuming to formulate an optimization problem involving PBEs. Thus, the moment method leads to a reduced-order model given by Equations (15-24).

Separate moment equations are used for the seed and nuclei classes of crystals, and they are defined as :

$$\mu_i^n = \int_0^{r_g} r^i n(r, t) dr \quad (9)$$

$$\mu_i^s = \int_{r_g}^\infty r^i n(r, t) dr \quad (10)$$

n in the superscript represents the nucleated crystal whereas **s** stands for the seeded crystal, r_g gives the critical radius separating the two. The moment equations for nucleated and seeded crystals are obtained as follows[?] :

Table 1. Parameter Values for Batch Crystallizer[? ?]

Parameters	Experimental Values
Growth Kinetics	
k_g	$1.44 \times 10^8 \mu m s^{-1}$
E_g/R	4859K
g	1.5
Nucleation Kinetics	
k_b	$285(s\mu m^3)^{-1}$
E_b/R	7517K
b	1.45

1. Nucleated crystals[? ?]

$$\frac{d\mu_0^n}{dt} = B(t) \quad (11)$$

$$\frac{d\mu_i^n}{dt} = iG(t)u_{i-1}^n(t) \quad i = 1, 2, 3 \quad (12)$$

2. Seeded crystals[? ?]

$$\mu_0^s = \text{constant} \quad (13)$$

$$\frac{d\mu_i^s}{dt} = iG(t)u_{i-1}^s(t) \quad i = 1, 2, 3 \quad (14)$$

The total moment is obtained as the summation $\mu_i^t = \mu_i^n + \mu_i^s$. Here, μ_0^s represents the amount of seeded crystals and hence remains constant. Also, fourth-order moments and higher do not affect third-order moments and lower, implying that only the first four moments and concentration can adequately represent the crystallization dynamics[?].

3. Deterministic Optimal Control

3.1. Case Study : Seeded batch crystallization

Optimal control involves the evaluation of time-dependent operating profiles, in terms of the control variable to optimize the process performance. In the crystallization domain, temperature becomes the control variable while the product yield is used to evaluate the performance of the model. The Optimization problem here is solved by using the method of Maximum Principle as discussed extensively by Diwekar[?]. This method stands out in compared to other techniques such as dynamic programming as it involves the use of first order ODEs while the later utilises partial differential equations. Another advantage of maximum principle is its ability to adapt to stochastic calculus which is explained in the next section.

In a seeded crystallization process it is essential to keep the nucleation phenomena to minimum as in the early stages of growth, nucleated crystals might compete with the seeded ones for growth. This ensures uniformity in the shape and size of the final product. It is achieved by incorporating the volume of nucleated in the **objective fuction**. Third moment(μ_3) represents volume in a crystallization model as evident from Equation(8) and is used as follows :

$$\max_{T(t)} \{ \mu_3^s(t_f) - \mu_3^n(t_f) \} \quad (15)$$

The active constraints for the process are given by :

$$C_s \leq C \leq C_m \quad (16)$$

C_m is the metastable concentration described by Equation(6). The state variables y_i for the process are now representes as :

$$y_i = \left[\begin{array}{cccccccc} C & \mu_0^s & \mu_1^s & \mu_2^s & \mu_3^s & \mu_0^n & \mu_1^n & \mu_2^n & \mu_3^n \end{array} \right] \quad (17)$$

Using the above notations the state equations can be realized as [?] :

$$\frac{dy_1}{dt} = -3\rho k_v G(t)(y_4 + y_8) \quad (18)$$

$$\frac{dy_2}{dt} = 0 \quad (19)$$

$$\frac{dy_3}{dt} = G(t)y_2 \quad (20)$$

$$\frac{dy_4}{dt} = 2G(t)y_3 \quad (21)$$

$$\frac{dy_5}{dt} = 3G(t)y_4 \quad (22)$$

$$\frac{dy_6}{dt} = B(t) \quad (23)$$

$$\frac{dy_7}{dt} = G(t)y_6 \quad (24)$$

$$\frac{dy_8}{dt} = 2G(t)y_7 \quad (25)$$

$$\frac{dy_9}{dt} = 3G(t)y_8 \quad (26)$$

$$(27)$$

Thus, the complete model involving the moment equations consists of nine state equations.

3.2. Solution technique

The algorithm of Steepest Ascent utilizes the maximum principle using the Hamiltonian to move towards the optimum value of temperature and maximise the objective function. To formulate the equations an adjoint variable (z_i) is defined corresponding to each state variable (y_i). The adjoint variable satisfies the relations given by eq 31. Using the state and the adjoint equations, a two-point boundary-value-problem is constructed. Here, the initial values for the states and the final conditions for the adjoint variables are known beforehand. The Hamiltonian can now be defined by eq 28.

$$H = \sum_{i=1}^9 z_i f(y_i, t, T) \quad (28)$$

The complete model with the objective function is given by Equations (29) to (31):

$$\max_{T(t)} \{y_5(t_f) - y_9(t_f)\} \quad (29)$$

$$\frac{dy_i}{dt} = f(y_i, t, T) \quad (30)$$

$$\frac{dz_i}{dt} = \sum_{j=1}^9 z_j \frac{\partial f(y_j, t, T)}{\partial y_i} = f(y_i, z_i, t, T) \quad (31)$$

with the initial and final conditions:

$$t_0 = 0 \quad \text{and} \quad t_f = 1800s \quad (\text{batch time})$$

$$y_i(t_0) = [0.1743 \quad 66.66 \quad 1.83 \times 10^4 \quad 5.05 \times 10^6 \quad 1.93 \times 10^9 \quad 0.867 \quad 0 \quad 0 \quad 0]$$

$$z_i(t_f) = [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad -1]$$

The steps followed to solve the optimal control problem have been discussed in detail below.

- An initial temperature $T(t) = 323K$ is assumed for the entire time horizon.

- The differential equations for state variables are integrated in the forward direction using the given initial conditions.
- The values of the adjoint variables are computed by backward integration of the adjoint equations given by eqs. 31.
- For evaluation of the Hamiltonian derivative, an analytical method proposed by Benavides and Diwekar[?], is used in which an additional variable corresponding to each of the state and adjoint variable is introduced. The variable θ_i corresponds to each of the state variable y_i and the variable ϕ_i corresponds to each of the adjoint variable z_i , which is given by eq(32).

$$\theta_i = \frac{dy_i}{dT} \quad \text{and} \quad \phi_i = \frac{dz_i}{dT} \quad (32)$$

- The Hamiltonian derivative is now calculated at each time step by the eq(33)

$$\frac{dH}{dT} = \sum_{i=1}^9 \left(\frac{dH}{dy_i} \right) \left(\frac{dy_i}{dT} \right) + \sum_{i=1}^9 \left(\frac{dH}{dz_i} \right) \left(\frac{dz_i}{dT} \right) \quad (33)$$

- The computed value of the derivative is checked against the convergence criterion ($\frac{dH}{dT} < \text{tolerance}$). If it is not satisfied, the temperature $T(t)$ is updated using the relation [?] given by eq(34). This ensures that the optimal control variable $T(t)$ is obtained using the extremum of hamiltonian.

$$T^{new}(t) = T^{old}(t) + M \left(\frac{dH}{dT} \right) \quad (34)$$

- The concentration is evaluated at that time step and compared with first with the saturation concentration (C_s) and then the metastable concentration(C_m) to validate the active constraints. If it found to be lesser than saturation value or greater than the metastable value, eqs (5) and (6) are used to compute the new temperature respectively.
- Iterations of above steps are repeated until the convergence criteria is achieved.

3.3. Results

The integration of differential equations was performed in python using the odeint routine from the scipy open source package. The resultant optimal temperature profile is show in Figure (3.3). The system attains a constant value of 300K at $\sim 1700s$. The decreasing temperature profile is indicative of the cooling crystallization. Also, the objective function can be observed to reach a maximum value in the Figure(3.3). The concentration profile of solute in the system is given by the Figure 3.3.

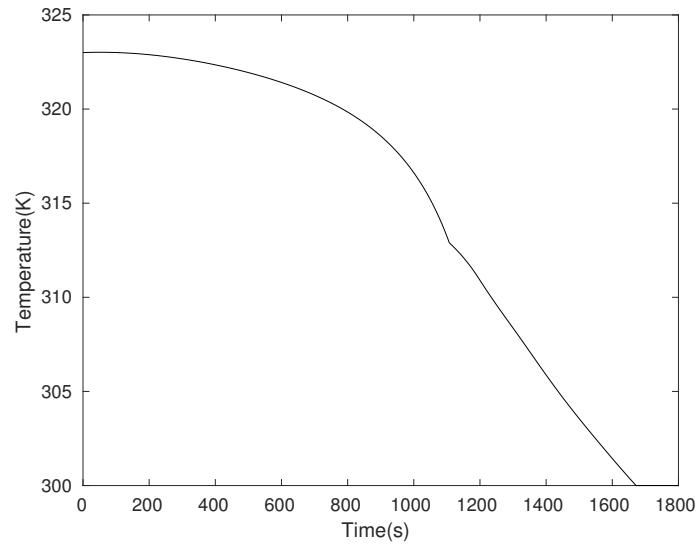


Figure 1. The cooling profile for the controlled variable $T(t)$ obtained at the final iteration

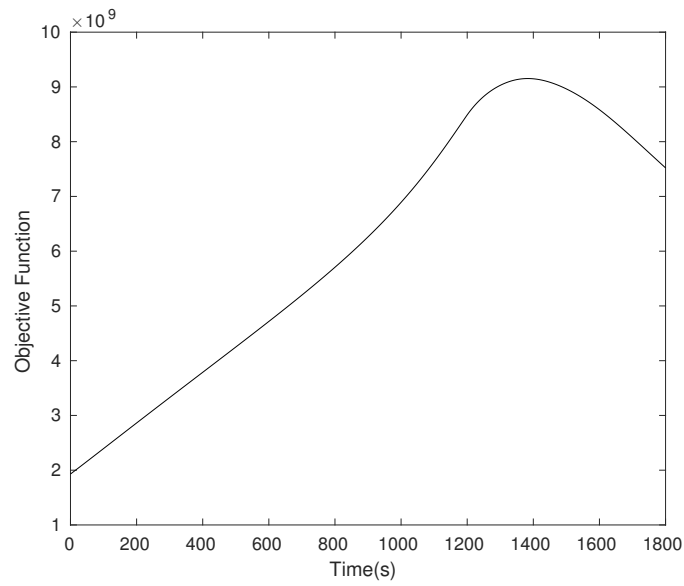


Figure 2. Objective Function $(\mu_3^s(t) - \mu_3^n(t))$

4. Optimal Control using Uncertainty Quantification

Kinetic parameters used in the Section(3) are empirical constants derived by fitting experimental data to the model, and hence, are a source of errors within the model. The exact values of the parameters are unknown and thus are termed as uncertainties. These errors tend to produce to sub-optimal results when used for real-world applications. In the

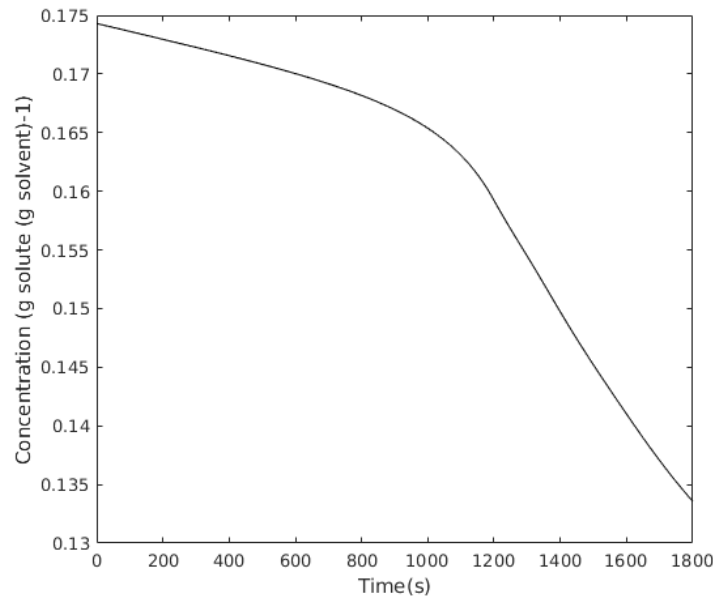


Figure 3. The Concentration profile as obtained.

next section, we convert the deterministic problem into a stochastic one wherein the kinetic constants are treated as random variables to include variation in their values into the final output produced by the model. Two methods are proposed to address the given problem.

4.1. Stochastic Optimal Control using Ito Processes

From several previous works([?]) it has been shown that the dynamic uncertainties present in batch reactors[?] and batch distillations[?], can be represented using stochastic processes called as the Ito processes. The growth and nucleation expressions have empirical constants shown in Table 1, which can be assumed to follow a Gaussian distribution[?]. We characterize the time-dependent uncertainties in the state variables using Ito processes known as **Brownian motion** with drift[? ?]. The advantage lies in the ability to integrate the equations through the principles of stochastic calculus and the use of stochastic maximum principle to solve for the optimal temperature profile. A simple Ito process can be written as eq (35)

$$dy = a(y, t)dt + b(y, t)dz \quad (35)$$

where dz is the increment of the Wiener process equal to $\varepsilon_t(\Delta t)^{1/2}$, and $a(y, t)$ and $b(y, t)$ are known functions. The random value ε_t has a unit normal distribution with zero mean and a standard deviation of 1. To estimate the values of the functions a and b , a generalized method presented by Diwekar[?] has been used.

In this work, equation(35) is used to modify the deterministic state (Equations (18) to (26)) to incorporate the

uncertainties into the moment equations as follows :

$$dy_1 = [-3\rho k_v G(t)(y_4 + y_8)] \Delta t + g_1 \varepsilon_1 \sqrt{\Delta t} \quad (36)$$

$$dy_2 = 0 \quad (37)$$

$$dy_3 = (G(t)y_2)\Delta t + g_3 \varepsilon_3 \sqrt{\Delta t} \quad (38)$$

$$dy_4 = (2G(t)y_3)\Delta t + g_4 \varepsilon_4 \sqrt{\Delta t} \quad (39)$$

$$dy_5 = (3G(t)y_4)\Delta t + g_5 \varepsilon_5 \sqrt{\Delta t} \quad (40)$$

$$dy_6 = (B(t))\Delta t + g_6 \varepsilon_6 \sqrt{\Delta t} \quad (41)$$

$$dy_7 = (G(t)y_6)\Delta t + g_7 \varepsilon_7 \sqrt{\Delta t} \quad (42)$$

$$dy_8 = (2G(t)y_7)\Delta t + g_8 \varepsilon_8 \sqrt{\Delta t} \quad (43)$$

$$dy_9 = (3G(t)y_8)\Delta t + g_9 \varepsilon_9 \sqrt{\Delta t} \quad (44)$$

$$(45)$$

$a(y, t)$ in each equation is replaced by the corresponding deterministic function for the state variable. Here, the g_i values represent the variance in the variable for which they are associated. They are calculated by recording the variance of the differences $(y_i^t - y_i^{t-1})$, which is divided by the time interval Δt , and then the square root of this value is taken. The **objective function** is also modified in Equation(46) for its evaluation in the stochastic domain.

$$\max_T L = \mathbf{E} [\mu_3^s(t_f) - \mu_3^n(t_f)] \quad (46)$$

\mathbf{E} is the expected value of the variable. The **Active Constraints** and **Initial Conditions** remain the same as mentioned in Section (3).

4.1.1. Solution Technique

The Hamiltonian for this section is modified to incorporate the uncertainties as[?] :

$$H = \sum_{i=1}^9 \left(z_i f_i + \omega_i \frac{g_{y_i}^2}{2} \right) \quad (47)$$

f_i are the deterministic parts for the eq (5.2-5.11). ω_i is an additional adjoint variable defined to calculate the Hamiltonian for the **Stochastic Maximum Principle** formulation[?].

The **Algorithm** for the method remains same as mentioned in Section (4.1.1) with minor changes.

1. Steps (1-4) are repeated.
2. The variable i corresponds to each of the state variable y_i and the variable i corresponds to each of the adjoint variable z_i , i corresponding to each i respectively.
3. The variable θ_i corresponds to each of the state variable y_i and the variable ϕ_i corresponds to each of the adjoint variable z_i , ψ_i for each ω_i respectively.
4. The Hamiltonian derivative is now calculated at each time step as :

$$\theta = \frac{dy_i}{dT} \quad \phi_i = \frac{dz_i}{dT} \quad \psi = \frac{d\omega_i}{dT} \quad (48)$$

$$\frac{dH}{dT} = \sum_{i=1}^9 \left(\frac{dH}{dy_i} \right) \left(\frac{dy_i}{dT} \right) + \sum_{i=1}^9 \left(\frac{dH}{dz_i} \right) \left(\frac{dz_i}{dT} \right) + \sum_{i=1}^9 \left(\frac{dH}{d\omega_i} \right) \left(\frac{d\omega_i}{dT} \right) \quad (49)$$

5. The convergence criteria and the constraints remain same as the above referenced method.

5. References

[1] ...