

Model Predictive Control: A Unified Approach for Urea-Based SCR Systems

2010-01-1184

Published
04/12/2010Thomas L. McKinley
Cummins, Inc.Andrew G. Alleyne
Univ. of Illinois at Urbana-Champaign

Copyright © 2010 SAE International

ABSTRACT

Despite the fact that urea-based selective catalytic reduction (SCR) of NO_x is a key technology for achieving on- and off-highway diesel emission standards, significant control challenges remain. Transient operation, combined with dramatic changes in catalyst dynamics over the operating range, cause highly nonlinear system behavior. Moreover, these effects depend on catalyst formulation and new catalysts continue to be developed. With many controllers, any difference in catalyst formulation, converter size, and engine emissions calibration require control system re-tuning.

To minimize control development effort, this paper presents a novel “generic” controller for SCR systems. Control action is grounded in a physics-based, nonlinear, embedded model. Through the model, controller parameters are adjusted a priori for catalyst formulation and converter size. The few remaining tuning levers are quite intuitive, and require no special knowledge of controls theory. The controller design is a type of model predictive control, and includes novel features for on-line determination of simplified, linear models as well as urea dosing limits for NH₃ slip control. Robust performance is achieved using feedback from NO_x and NH₃ sensors. Controller capability is demonstrated for SCR inlet conditions corresponding to an EPA 2007 heavy-duty diesel operating on the European Steady-State Cycle (ESC) and US EPA heavy-duty FTP cycle.

INTRODUCTION

Due to their durability and fuel economy, diesel engines are the power-plant of choice for trucks, buses, agricultural tractors, and earth-moving equipment. Unfortunately,

significant amounts of NO and NO₂ are formed during the combustion process. The health effects of these NO_x emissions have led nations in Europe, Asia, and North America to implement aggressive emissions limits for both on- and off-road diesel-powered vehicles. In some cases, the standards are so stringent that they cannot be attained by production in-cylinder methods. Instead, some type of NO_x aftertreatment is required. After considering several different approaches, Johnson [1] showed that the technology of choice for many applications is selective catalytic reduction (SCR).

Moreover, SCR system performance requirements are themselves quite challenging. For example, the 2007 US on-road heavy-duty diesel emission standards limited NO_x emissions to 1.61 g/kW-hr. To achieve compliance, all manufacturers chose to use only in-cylinder NO_x emissions control. In 2010, the emissions limit was reduced to 0.27 g/kW-hr. All manufacturers complying with this standard will do so using SCR. In particular, to meet the new standard with minimal engine design change, a NO_x conversion efficiency of 80% is required. In pursuing this goal, tailpipe NH₃ mole fraction, commonly called NH₃ slip, is limited to 10 ppm under steady-state conditions and 20 ppm during transients. Although NH₃ emissions are unregulated, these self-imposed limits are comparable to occupational health standards and the detection threshold of ammonia odor by the human nose [2,3].

To meet these performance requirements, the injection rate of the reductant (either a urea-water solution or gaseous NH₃) must be carefully managed. Failure to do so results in excessive NO_x emissions or NH₃ slip. Furthermore, the commanded injection rate must take into account the

dynamics of NH_3 adsorption and desorption. This can be particularly challenging for applications characterized by wide ranging and highly transient operating conditions. Clearly, reductant injection control is a critical factor affecting system performance, and so many different SCR controller designs have been proposed [4,5,6,7,8,9,10,11].

In pursuit of optimal performance, these controllers frequently use maps for NO_x conversion efficiency or NH_3 storage [4,5,6,9,10,11] or switched controllers [9,10,11] to take into account NH_3 slip limitations. Although these features do provide performance improvement, they increase time and effort associated with controller tuning. In addition, each time the SCR converter or engine calibration is changed, the tuning effort may need to be repeated. Since manufacturers frequently produce a family of engines having different displacements and ratings, and the calibration effort is repeated at least twice during the development program, the cost of controls tuning is multiplied several times over.

As an alternative, this paper presents a “generic” SCR controller in which reductant injection rate is determined from a physics-based model and a few intuitive tuning parameters. This map-less controller makes a priori controller adjustments in response to changes in catalyst formulation, converter size, and SCR inlet conditions. A type of model predictive controller, this approach greatly simplifies the tuning effort, while still providing excellent system performance.

SCR SYSTEM DESCRIPTION

Figure 1 is a schematic of a complete diesel exhaust gas aftertreatment system as studied in this paper. The diesel oxidation catalyst (DOC) and diesel particulate filter (DPF) reduce hydrocarbon and particulate matter emissions. Platinum-based catalytic coatings on these components promote the oxidation of NO to NO_2 , which is advantageous both for oxidizing the soot trapped by the DPF and improving low temperature NO_x conversion efficiency in the SCR system. The NO_x mole fraction entering the aftertreatment system is measured using a NO_x sensor. Although this portion of the system has no effect on total NO_x emissions, it alters the gas temperature and molar ratio of NO_2 to NO_x entering the downstream SCR system.

The principle components of the SCR system are as follows. A doser injects a urea-water solution, commonly called diesel exhaust fluid (DEF), in response to the controller's command. The NH_3 required to reduce NO_x is supplied by decomposition of the urea contained in DEF. The gas temperature is measured upstream and downstream of the converter. The converter itself consists of 35% porosity cordierite substrates with 400 cells per square inch and a wall thickness of 7 mil. A copper-zeolite catalyst is coated onto the substrate surfaces. An NH_3 sensor and a second NO_x

sensor are placed downstream of the converter. This sensor arrangement addresses known issues with NH_3 cross-sensitivity of current day NO_x sensors. The controller is implemented on an electronic control module (ECM), which also has access to engine data such as exhaust mass flow rate and air-fuel ratio over the controller area network (CAN).

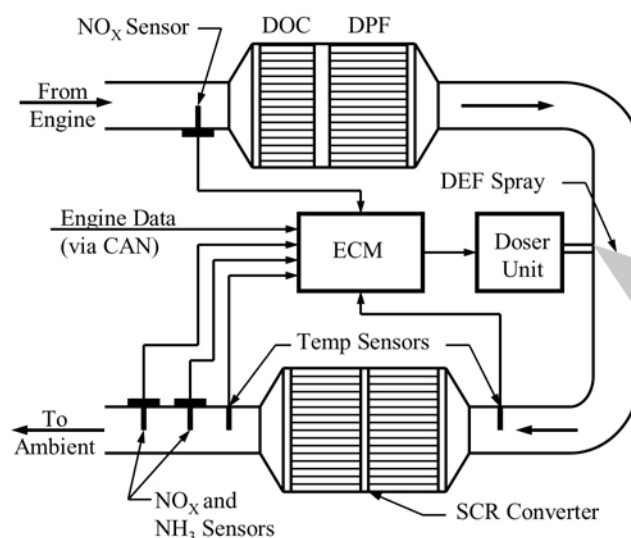


Figure 1. Diesel Exhaust Aftertreatment System

MODEL PREDICTIVE CONTROL PARADIGM

A model predictive controller contains three essential features [12,13]: (a) a model to predict the output over a future output horizon; (b) control action, over a future control horizon, chosen to minimize some cost function; and (c) a receding horizon strategy whereby, at each sample time, only the first control is implemented and the optimization is repeated at the next sample time. The control and output horizons are illustrated in Figure 2 for the k th sample time. The input and the output are updated using the same sample interval T .

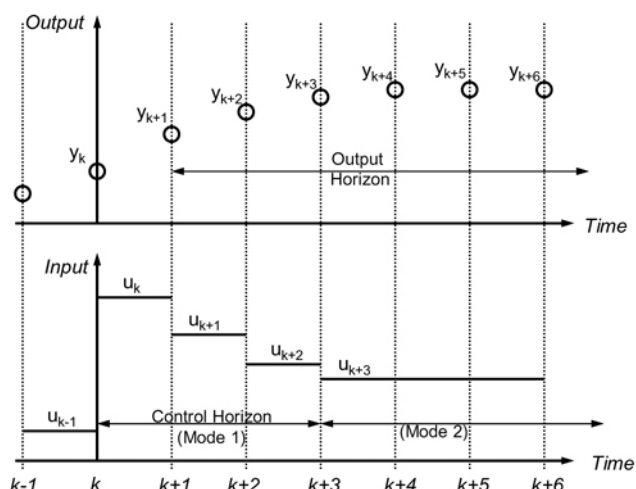


Figure 2. Control and Output Horizon Nomenclature

Model predictive control (MPC) originated in the chemical industry [14], but has spread to many different fields. In just the last seven years, 13 automotive applications have been published, including traction control systems, air-handling systems, HCCI engine valve trains, fuel injectors, lean NOx traps, and three-way catalysts. However, this paper is the first known MPC application to an automotive SCR system.

MPC's advantages and disadvantages are as follows. Its primary advantage is that it takes constraints directly into account when determining control action. This is especially relevant for SCR systems because of the NH₃ slip limits mentioned above. On the other hand, its two primary disadvantages stem from the need to solve an optimization problem at each sample time. These are: (a) computational efficiency; and (b) solution feasibility. Specifically, the time required to solve the optimization problem introduces a delay in the controller. This is undesirable and can become problematic if it is longer than the sample interval. Moreover, if there is no feasible solution (meaning, one that meets all constraints), which control to apply becomes problematic.

MPC can be best described as a paradigm rather than a specific control technique because it accommodates so many different types of models, optimizers, methods of including constraints, and control law formulations. These features must be chosen carefully since they can significantly affect computational efficiency and solution feasibility, and can be the difference between practical and impractical controllers.

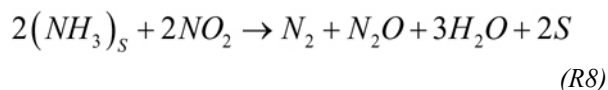
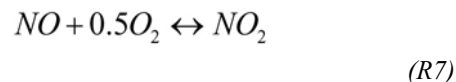
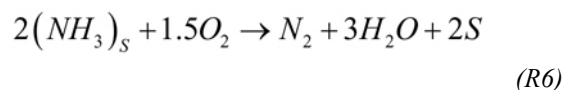
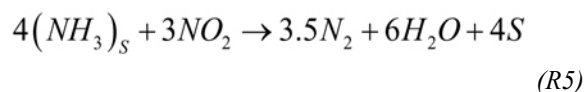
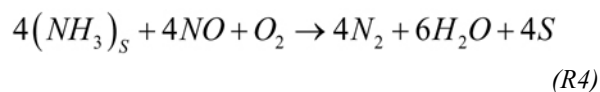
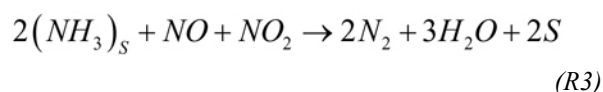
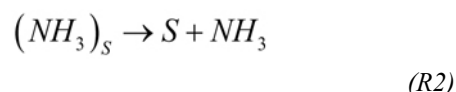
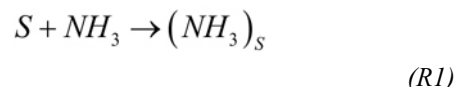
SCR SYSTEM EMBEDDED MODEL

SCR system performance is frequently predicted using a one-dimensional model, where the converter is divided into a number of axial elements. The specific model used in this work is described elsewhere [15], but in brief each element has two states: (a) catalyst temperature; and (b) NH₃

coverage. Due to its short residence time, the gas phase is modeled using quasi-steady equations.

CHEMICAL REACTIONS

The model includes the following chemical reactions:



In these equations, NH_3 refers to gaseous ammonia, $(NH_3)_S$ refers to adsorbed ammonia, and S refers to an unoccupied catalytic site.

The reaction mechanism given above includes nearly all significant effects occurring on SCR catalysts. By adjusting activation energies and frequency factors, it is possible to represent copper-zeolite, iron-zeolite, and vanadium-pentoxide based formulations. In this respect, the model can be considered a "generic" SCR converter model.

As stated in the previous section, the NH₃ required to reduce NOx is formed from the urea contained in DEF. Urea decomposition is a fairly complicated process and continues to be an active area of research in SCR systems [16]. Instead, the NH₃ concentration entering the converter is computed by assuming instantaneous and complete conversion of urea to NH₃. Although this is not a good assumption based on

experimental measurements [5], this is a very common approach in SCR system models. Moreover, several researchers have demonstrated good model accuracy even with this simplification [6, 8].

MODEL SPATIAL RESOLUTION

For a typical SCR converter, a total of 24 uniform length elements are required to accurately resolve spatial distributions in coverage [17]. Although this still permits real-time simulation on laptop computers, run-time and memory requirements are excessive for use as an embedded model within the ECM. Thus, the relationship between model accuracy and the number of elements is a central issue for any type of model based control scheme, including MPC. Prior work on this subject resulted in recommendations ranging from one or two elements [6,7,8] to as many as 8-12 elements [11].

Following the work of Herman et al [11], model resolution was evaluated using the ESC test. In particular, Figure 3 shows the error in mass average NOx conversion efficiency and maximum NH3 slip for models of varying resolution. The results for a 24 element model are considered to be the true values, and gave a mass average NOx conversion efficiency of 0.855 and a maximum slip of 6.3 ppm. The commanded molar ammonia-to-NOx ratio (ANR) was the same in each case, as were the model parameters. Using error limits of 10 ppm for NH3 slip and 0.02 for NOx efficiency indicates that at least six elements are needed.

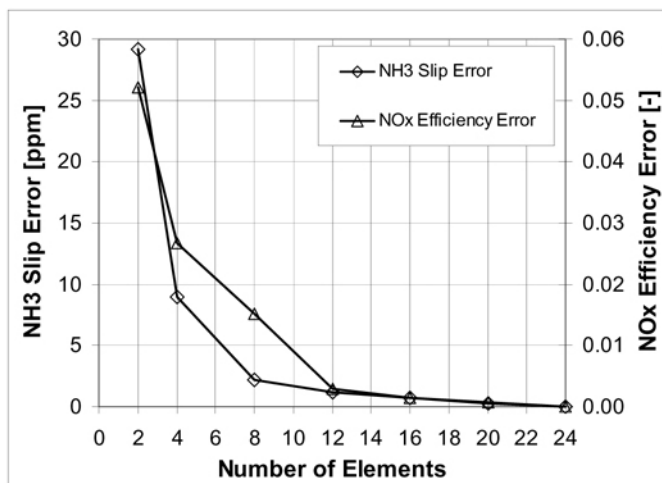


Figure 3. Error in Maximum NH3 Slip and Mass Average NOx Conversion Efficiency on ESC Test with Varying Number of Elements

However, recalibrating model parameters as resolution is changed is an effective way of reducing model error. Using that approach, the error of a two element model was lowered to 0.018 in NOx efficiency and 7.6 ppm in NH3 slip, which meets the error criteria stated above. This is a substantial

improvement from the errors of 0.052 and 29.2 ppm without recalibration. Thus, for the remainder of this work a two element embedded model is employed.

Using two elements includes spatial distribution of coverage and temperature, if only to a limited degree. These effects are known to be significant, and that is a distinct advantage of a two element model compared to a one element model [7,8]. Furthermore, a 12 element model has been shown to be compatible with ECM run-time and memory requirements [11]. Hence, a two element model would also be ECM compatible.

APPLICATION OF MPC TO SCR

Having provided a brief overview of MPC and introduced the embedded model, the application of MPC to SCR system control is now described.

CONTROLLED INPUT

In the arrangement shown in Figure 1, SCR system performance is controlled by adjusting the DEF injection rate, which in turn determines the amount of NH3 supplied to the converter. Once NH3 is adsorbed onto the catalyst surface, several pathways for NOx reduction are possible. However, all of them (reactions (R3), (R4), (R5), and (R8)) are expressed in terms of a molar ratio of NH3 to NO and NO2. Therefore, the molar ratio of supplied NH3 to supplied NOx is taken to be the system's controlled input. Usually called the ammonia-to-NOx ratio or ANR, this non-dimensional variable is signified by the symbol u in accordance with controls literature convention.

CONTROLLED OUTPUT AND REFERENCE

The SCR system's function is to reduce the amount of NOx in the exhaust stream. For that reason, the controlled output is taken to be the NOx conversion inefficiency. Defined as the NOx mole fraction at the converter outlet divided by the NOx mole fraction at the converter inlet, this quantity is also non-dimensional. Again, in agreement with the controls literature, it is represented by the symbol y .

In closed loop control, the input (u) is varied to reduce error between the output (y) and a set-point or reference value, which in this case is signified by the symbol y_{REF} .

COST FUNCTION

Referring to Figure 2, the cost function to be minimized is:

$$J = \sum_{i=1}^{ny} (y_{REF} - y_{k+i})^2 + \lambda \sum_{i=0}^{nu-1} (\Delta u_{k+i})^2 \quad (1)$$

The first summation on the right-hand side is the output tracking error, while the second one is the control effort. The trade-off between these two terms is adjusted through the weighting factor λ . Note that control effort is in terms of the input increment, which is defined as:

$$\Delta u_{k+i} = u_{k+i} - u_{k+i-1} \quad (2)$$

The sampling interval is taken to be one second, in agreement with the capabilities of the DEF doser.

The length of the output horizon ny is taken to be infinite, as recommended in the literature [12]. This means the input must also be defined over an infinite horizon. To limit the degrees of freedom in the optimization, it is assumed to vary only over the first nu time steps and is then held constant, an approach that is commonly called *no terminal control*. As recommended in the literature [12], nu is taken to be 3. Thereafter, the input is held constant at the value giving $y = y_{REF}$.

SENSOR FEEDBACK

Some mismatch can always be expected between the plant and model. However, the controlled output at the current time k can be measured using the inlet and outlet NOx sensors. This allows the model error to be calculated from:

$$d = y_{MEAS,k} - y_{PRED,k} \quad (3)$$

Here, the subscripts MEAS and PRED refer to sensor measurements and model predictions, respectively. This error can be considered an output disturbance, which is assumed to be constant over the output horizon. The values of the output in the cost function are then computed from:

$$y_{k+i} = y_{PRED,k+i} + d \quad (4)$$

By combining equations (3) and (4), it can be seen that this procedure zeroes out the model error at the start of the output horizon ($i = 0$). This same procedure is applied to NH3 slip predictions.

Since model error changes over time, the output disturbance in equation (3) is updated each sample interval. This periodic

correction for model error is the usual manner of including sensor feedback into MPC, and converts it from open-loop optimal control to closed loop control [18].

ON-LINE FEASIBILITY CHECK

Feasibility, meaning the existence of control action meeting constraints, is a fundamental issue in MPC. In this case, there are two constraints:

$$0 \leq u \leq u_{MAX} \quad (5)$$

$$0 \leq z \leq z_{MAX} \quad (6)$$

Here the variable z is the NH3 slip in ppm. The primary concern is that constraints may preclude reaching the desired NOx conversion inefficiency y_{REF} . To simplify the analysis in the face of uncertain future events, we assume catalyst temperature and inlet conditions do not change over the output horizon. Then, for any u , the corresponding equilibrium values of coverage and the outputs y and z can be determined from the embedded model (including disturbances). The input is then iterated until $y = y_{REF}$ or the constraints in equations (5) and (6) are satisfied. If constraints need to be violated to reach $y = y_{REF}$, the control problem is infeasible. This is resolved by reassigning y_{REF} to the limiting value consistent with constraints.

This approach greatly simplifies control tuning. Specifically, only a single value of y_{REF} is given. If this target exceeds the achievable NOx conversion inefficiency because of NH3 slip or doser capacity limitations, then the target is adjusted automatically. The adjustment includes NOx and NH3 sensor measurements and is not strictly limited by model accuracy. This can be envisioned as on-line determination of NOx conversion efficiency or coverage maps, freeing the engineer of that task.

PATH TO COMPUTATIONAL EFFICIENCY

As described above, the MPC controller solves a constrained optimization problem, minimizing the cost function in equation (1) subject to the constraints in equations (5) and (6). The problem has nu degrees of freedom, which are the components of the input increment schedule during the control horizon. This schedule can be written in vector form as:

$$[\Delta u] = [\Delta u_k \quad \Delta u_{k+1} \quad \cdots \quad \Delta u_{k+nu-1}]^T \quad (7)$$

However, the optimization problem still appears daunting. In particular, the embedded model is nonlinear and can only be integrated numerically. Applying it over an infinite output horizon for even one input increment schedule would take an infinite amount of computer time.

Clearly, such an approach is impractical. Instead, simplified prediction models, analytical solutions, and cost function reformulation are used to recast the optimization problem into a manageable framework.

SIMPLIFIED PREDICTION MODELS

The embedded model has four states, which are the catalyst temperature and NH₃ coverage for each of the two elements. In the face of uncertain future events, it can be assumed that catalyst temperature and SCR converter inlet gas conditions are constant over the output horizon (with the exception of the controlled input of ANR). This reduces model order from four to two. A further simplification is to represent the coverage by its average value rather than the coverage of each element. This reduces model order from two to only one.

For small perturbations in ANR, it can be shown that average coverage is represented quite well by a simple first order system having a time constant and a steady-state gain [17]. Moreover, since NO_x reduction is accomplished using adsorbed rather than gaseous ammonia, for small perturbations the NO_x conversion inefficiency can be approximated as linear function of average coverage. Therefore, the fourth order, nonlinear embedded model can be approximated using the first order, linear model:

$$\frac{d}{dt}(\bar{\theta} - \bar{\theta}_0) = -\frac{1}{\tau_\theta}(\bar{\theta} - \bar{\theta}_0) + \frac{K_\theta}{\tau_\theta}(u - u_0) \quad (8)$$

$$y = y_0 + \left(\frac{dy}{d\bar{\theta}} \right) (\bar{\theta} - \bar{\theta}_0) \quad (9)$$

In the equations above, the subscript 0 refers to the point about which the model is linearized.

A 24 element plant simulation was then applied to understand how some of the linear model parameters change over the engine operating range. The simulation was applied for 14 different steady-state conditions, with temperature ranging from 193 to 518 deg C and space velocity ranging from 7.9 to 61.5 K/hr. ANR was adjusted until either a desired NO_x

conversion inefficiency was attained or the NH₃ slip reached 10 ppm, whichever gave the lower value of ANR. ANR was then perturbed and the time constant was found from a least squares curve fit of the simulated response.

Figure 4 shows that, for a given NO_x conversion inefficiency, there is an exponential relationship between average coverage and catalyst temperature. Although NO_x conversion inefficiency was varied in uniform increments, this did not give uniform increments in coverage. That indicates the derivative of NO_x inefficiency with coverage is a function of both temperature and coverage, a fact verified by more extensive simulation [17]. Figure 5 shows that the time constant varies exponentially with temperature but is also sensitive to coverage. Although not shown, similar sensitivities were found for the steady-state gain.

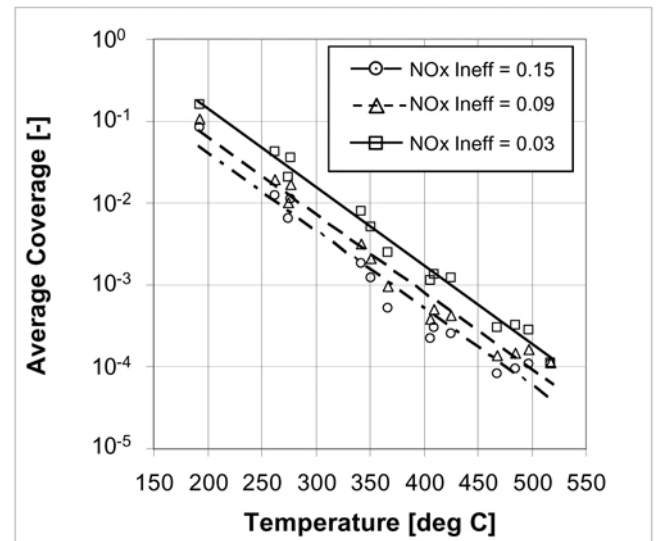


Figure 4. Average Coverage for Different NO_x Inefficiencies as a Function of Catalyst Temperature