

# Sustainable urban transport development: a modelling approach

M. H. P. Zuidgeest & M. F. A. M. van Maarseveen

*Centre for Transport Studies, University of Twente, The Netherlands*

## Abstract

Current transport systems and transport planning methods and models are not necessarily compatible with the requirements of sustainable transport development. Adequate transport systems can only be obtained by use of a sustainable transport paradigm and an accompanying analytical framework. Therefore, this paper presents a theoretical framework, which is based on a paradigm for sustainable transport development. This paradigm advocates a comprehensive decision-making that anticipates and manages scarce resource use, including environment and finance, while developing the transport system in terms of quality of access and/or person throughput. Furthermore, a simplified version of a dynamic optimisation model that can assist in the complex and political decision-making process with respect to sustainable transport development is introduced, based on the conceptualization and characterization of the sustainable urban transport development problem as a constrained optimisation problem. Based on Pontryagin's Maximum Principle, the dynamic model reveals control paths for achieving a sustainable and developed transport system. The model in its present form can be applied directly to strategic networks of limited numbers of (aggregated) zones and (aggregated) links.

*Keywords: sustainable development, urban transport, dynamical modelling.*

## 1 Introduction

Mobility of people (and freight) is an essential prerequisite for social - economic development. In most cities throughout the developed and developing world, however, motorised vehicles, notably cars and trucks, have become the most important means of mobility, at the cost of non-motorised transport as well as public transport.



Because of this motorisation, congestion, traffic unsafety, air and noise pollution, changing land-use patterns, social isolation etceteras, have become common and widespread images in cities. The usual solution to these emerging problems and changing requirements in transport is to build extra capacity, make better use of existing infrastructure, discourage and/or promote other means of transport or even influence travel patterns of people as well as freight, following the principle of *predict-provide-manage*. Decisions to do so are supported by best-practices, theories, and tools of transport planning. However, by doing so, current transport systems and transport planning methods and models (used in developed as well as developing countries) are not necessarily compatible with the requirements of sustainable transport development.

In this paper a sustainable and developed urban transport system is postulated to be: 'a transport system that meets the people's transport related needs in terms of mobility, accessibility and safety, within limits of available or affordable environmental, financial and social resource capacities' as defined in Zuidgeest [1], following the Brundtland definition in [2].

This definition is based on the basic idea of sustainable development and its taxonomy that characterises sustainable development as consisting of two distinct elements, that is sustainability as well as development, while having three different dimensions, that is economic and financial sustainability, environmental and ecological sustainability as well as social sustainability, furthermore having an analytical, a normative as well as a strategic level of discourse. Such taxonomy is well described in Gudmundsson and Höjer [3].

On the basis of this, a framework for sustainable urban transport development can be constructed that adopts a multi-directional conception of sustainable transport development, which tries to answer two important main questions, i.e.:

1. How can basic mobility and accessibility options to people be sustained or enhanced? [the development question];
2. How can limited transport related resources, that is environmental, social and economic resources capacities, be used to guarantee intergenerational equity? [the sustainability question].

In addition, sustainable systems, like ecological systems, are not steady-state systems, but rather dynamic systems with many feedback loops to provide self-regulation and to keep growth of each part of the system coordinated with the other parts as the system evolves, as stated by Replogle [4]. Similarly, Brundtland [2] defines it as 'sustainable development is a process, not a fixed state'. Therefore sustainable transport development as compared to sustainable development should also be regarded as a process, being intrinsically dynamic, as it is the coevolution of travel demand and infrastructure supply (and management) that makes a sustainable transport development.

Adopting this framework, transport professionals should be able to aim directly at reaching certain transport development objectives through their transport policies and plans, while maintaining non-declining levels of transport system performance as well as keeping resource-use levels below those maximally acceptable, in other words affordable or available, levels.



To implement this in transport planning and modelling practice, four requirements can be derived and should be implemented, see also [1].

The first requirement relates to policy objectives. Transport system performance (over time), which is the quality of functioning of the transport system at a given level of travel demand and infrastructure supply, should be explicitly related to the political direction chosen, which is accordingly translated into a (quantifiable) transport policy objective and implemented directly into the modelling. These objectives of transport system performance may vary from a mere motorised traffic orientation, in terms of level-of-service, to social indicators, in terms of equity, or even (weighted) combinations of objectives.

The second requirement deals with transport dynamics. Travel behavioural rules and transport system mechanisms have been studied in transport science, and implemented in transport models, extensively. A distinction is usually made in mechanisms of travel demand and infrastructure supply, which are considered to be in equilibrium. Travel demand is directly related to the social-economic realities of people (in terms of utility of trip making) as well as the transport network accessibility available to the people (in terms of disutility of trip making). Hence, travel demand is elastic to changes in both the social-economic realities as well as accessibility, and should be implemented in the modelling as such. Furthermore, these different mechanisms operate at different time-scales. A change in route choice-behaviour might require a relatively short period to transpire, whereas a change in infrastructure capacity might take a longer period to be accomplished. Therefore, it seems unlikely that an equilibrium between demand and supply exists in a dynamic model. A disequilibrium formulation is thus considered to be more appropriate. Hence, a system dynamics approach, including lagged-adjustment, disequilibrium models, should be used.

The third requirement is about resource capacities. Quantitative targets, preferably related to international standards for resource-use, for example those of the WHO or the legally binding targets for greenhouse gas emissions of the Kyoto protocol, should be set and internalised in transport models. To do this within the framework derived in this research, available resource capacities should be known. In particular, deriving environmental capacities for a demarcated urban area can be very difficult. In addition, non-point source emission models should be applied to estimate pollution levels. Likewise, financial capacity and spending, but at the same time also revenues from, for example, road pricing, should be internalised in the modelling.

The fourth and last requirement is about policy measures. Transport planners have several transport policy measures available. A distinction can be made in demand-side measures (related to the affordability of travel options) as well as supply-side measures (related to the availability of travel options). To control transition paths of state variables, as guided by the transport policy objectives, these measures can be deployed (if necessary, bounded by constraints) in a prescribed sequence and timing. Transport modelling should be able to derive these paths.



Incorporating all these requirements, sustainable transport development planning is objective-driven and resource-bounded for both current and future generations, while being infrastructure supply-limited rather than demand-driven. Hence, a proactive approach that can be characterised as one of provide - (manage -) predict (hence 'prevent'), unlike the common principle of predict - provide (- manage), remains. The remainder of this paper discusses an abridged version of a dynamic optimisation model [1] that can assist in this process of planning sustainable transport development.

## 2 Model description

Based on the requirements put forward in previous paragraph a dynamic optimisation model can be defined that optimises a transport policy objective over time (time  $t_0$  until  $t_1$ , which may be years) given the dynamics in travel demand (link travel demand  $V_l(t)$ ) and infrastructure supply (effective link capacity  $C_l^e(t)$ ). The latter can be controlled by adding or distracting (continuous and bounded) capacity  $U_l^c(t)$ . Furthermore, total emissions  $E(t)$  are monitored and bounded by environmental capacity  $E^*$ . This model is demonstrated for a small network as displayed in figure 1, where, for simplicity of the example, the exogenous social-economic and land-use variables, trip production  $Q(t)$  and trip attraction  $X(t)$  are assumed to be stable over time. The dynamic optimisation problem is given below (eqn. (1)).

$$\begin{aligned} \min \int_{t_0}^{t_1} \sum_l & \left( \left( \frac{V_l(t)}{C_l^e(t)} - \delta_l \right)^2 + \alpha_7 U_l^c(t)^2 + \chi_1 \max(0, (E(t) - E^*))^2 \right) dt \\ \text{s.t. } \frac{dV_l(t)}{dt} &= \gamma_1 \left( \hat{V}_l(t) - V_l(t) \right), \quad \forall l, \\ \frac{dC_l^e(t)}{dt} &= U_l^c(t), \quad \forall l, \\ \frac{dE(t)}{dt} &= \gamma_1 \left( \hat{E}(t) - E(t) \right), \\ \text{and } 0 \leq U_l^c(t) &\leq U_l^{c\max}, \quad \forall l, \\ &\text{boundary conditions at } t_0 \text{ and } t_1. \end{aligned} \tag{1}$$

Before discussing the actual minimization problem, i.e. the transport policy objective, the three equations of motion for the main variables  $V_l(t)$ ,  $C_l^e(t)$  and  $E(t)$  will be described.

If travel demand is regarded to be elastic to changes in accessibility, the equilibrium level  $\hat{V}_l(t)$  should also include a measure of elasticity. This implies that the *potential* trip generation capacity of a zone needs to be known. Hence, starting with a disequilibrium travel demand model, expressing the changing



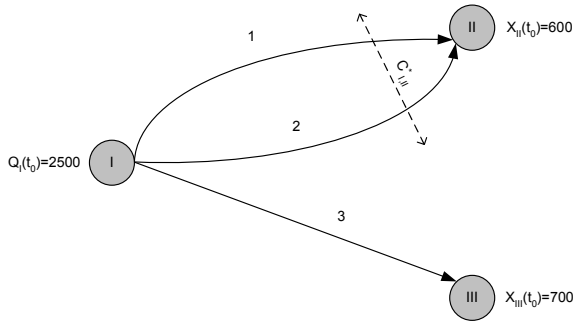


Figure 1: Test network with directed links  $l \in \mathcal{L} = \{1, 2, 3\}$ , trip production  $Q_I$ , and trip attractions  $X_{II}$  and  $X_{III}$  in  $[\text{pers h}^{-1}]$  at time  $t = t_0$ . Variable  $c_{I,II}^*$  indicates the composite costs for the corridor  $I - II$ .

state of travel demand (in person-car units  $[\text{pcuh}^{-1}]$ ) between the equilibrium traffic volume  $\hat{V}_l(t)$  and actual traffic volume  $V_l(t)$  is, following Donaghy and Schintler [5]:

$$\frac{dV_l(t)}{dt} = \gamma_1 \left( \hat{V}_l(t) - V_l(t) \right), \quad \forall l, \tag{2}$$

with  $\gamma_1$  the lagged-adjustment of travel demand to changes in accessibility. The equilibrium link demand function  $\hat{V}_l(t)$  is assumed to be built-up of:

1. the time-varying potential trip generation capacity of a zone  $i$ , based on social-economic characteristics of the zone,  $Q_i(t)$ ;
2. a demand elasticity factor  $D_{ij}(t)$ , representing the induced effect of accessibility between zones  $i$  and  $j$  on trip generation;
3. a simultaneous mode  $m$ , destination  $j$ , and route  $r$  choice model,  $G_{ijmr}(t)$ .

Hence, part  $D_{ij}(t)$  of the exogenous time-varying potential trip generation capacity  $Q_i(t)$  in zone  $i$  in  $[\text{pers h}^{-1}]$  is distributed over the different mode and route choice-options for the corridor  $(i, j)$ , using a discrete-choice model  $G_{ijmr}(t)$  and converted to vehicles in  $[\text{pcuh}^{-1}]$ , applying a vehicle occupancy factor  $\theta_{1m}$ , converting person-trips to vehicle-trips ( $[\text{pcuh}^{-1}]$ ):

$$\hat{V}_l(t) = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}^i} \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}_{ij}^l} \theta_{1m} (Q_i(t) D_{ij}(t) G_{ijmr}(t)), \quad \forall l, \tag{3}$$

with  $\mathcal{R}_{ij}^l$  the route serving origin-destination pairs that contain link  $l$ , or:  $\mathcal{R}_{ij}^l = \{r \in \mathcal{R}_{ij} \mid \exists n : r_n = l\}$ , and:  $r = (1, 2, \dots, l_r, \dots, L_r) \in \mathbb{N}^{N_r}$ , as well as  $\mathcal{J}^i$  the destination set excluding destination  $j = i$ , or:  $\mathcal{J}^i = \{j \in \mathcal{J} \mid \nexists n : j_n = i\}$ . Note that the set of origins  $i$  is:  $\mathcal{I} = \{1, 2, \dots, i, \dots, I\} \in \mathbb{N}^{N_i}$ , whereas the destinations  $j$  are:  $\mathcal{J} = \{1, 2, \dots, j, \dots, J\} \in \mathbb{N}^{N_j}$ . Furthermore, a mode-set is defined as:  $\mathcal{M} = \{1, 2, \dots, m, \dots, M\} \in \mathbb{N}^{N_m}$ .

The chance of simultaneously selecting a certain destination  $j$ , mode  $m$  and route  $r$  is obtained using a dynamised version of the well-known doubly-constrained gravity model, which is distributing trips on the basis of the utilities  $u_{ijmr}(t)$  for the different mode  $m$ , and route  $r$  choice-combinations as well as the exogenous time-varying trip attraction value  $X_j(t)$ :

$$G_{ijmr}(t) = \frac{X_j(t) \exp(-\lambda_1 u_{ijmr}(t))}{\sum_{j' \in \mathcal{J}^i} \sum_{m' \in \mathcal{M}} \sum_{r' \in \mathcal{R}_{ij'}} X_{j'}(t) \exp(-\lambda_1 u_{ij'm'r'}(t))}, \quad \forall i, j, m, r, \quad (4)$$

with scale parameter  $\lambda_1$ , and  $X_j(t)$  the destination attractiveness, or trip attraction.

The utility function  $u_{ijmr}(t)$  for generalised costs of travel is in this simplified example expressed as a linear equation of route travel time  $\tau_r(t)$  alone:

$$u_{ijmr}(t) = \beta_{3m} \sum_{l_r=1}^{N_r} \tau_{l_r}^0 \left[ 1.0 + \alpha_1 \left( \frac{V_{l_r}(t)}{C_{l_r}^e(t)} \right)^{\beta_1} \right], \quad \forall i, j, m, r, \quad (5)$$

with  $\beta_{3m}$  a mode-specific parameter (note that routes  $l_r$  are mode-specific as well and may use the same link). Parameter  $\beta_{3m}$  is the *value-of-time* parameter for conversion of time to monetary units. Route travel time is a summation of link travel times that comprise route  $r$ , which are expressed as strictly increasing, continuous and nonlinear functions of the volume  $V_i(t)$  to effective capacity  $C_{l_r}^e(t)$  ratio and free-flow travel time  $\tau_{l_r}^0$ , by applying the well-known BPR equation with parameters  $\alpha_1$  and  $\beta_1$ .

The maximum amount of revealed travel demand on a corridor (part of it may actually go to another destination  $j$ ), or origin-destination pair  $(i, j)$  is equal to:  $Q_i(t) \cdot D_{ij}(t)$ , where the elasticity factor  $D_{ij}(t)$  is formulated, based on De la Barra [6], as:

$$D_{ij}(t) = a_1 + b_1 \exp(-\lambda_2 (c_{ij}^*(t) - c_{ij}^0)), \quad \forall i, j. \quad (6)$$

Parameter  $a_1$  represents the ‘captive’ trips that will be performed irrespective of the composite costs and  $(a_1 + b_1)$  the maximum ratio, i.e. 100% of the trips that can be performed in ideal circumstances, or free-flow conditions.

The composite cost, or logsum cost  $c_{ij}^*(t)$ , on a origin-destination pair  $(i, j)$  are calculated by aggregating generalised costs (expressed in the utility function  $u_{ijmr}(t)$ ) over all modes  $m$  and routes  $r$  that serve the origin-destination pair:

$$c_{ij}^*(t) = -\frac{1}{\lambda_3} \ln \left[ \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}_{ij}} \exp(-\lambda_3 u_{ijmr}(t)) \right], \quad \forall i, j, \quad (7)$$

where  $\lambda_3$  is the scale parameter to obtain the expected maximum utility of the choice-set:  $\mathcal{C} = (\mathcal{R}_{ij}, \mathcal{M})$ .

The ‘free-flow’ composite costs  $c_{ij}^0$  can be derived by calculating the generalised costs for all choice-options under free-flow conditions.



The performance of a road link in terms of the volume-over-capacity ratio can be improved by expanding (or decreasing) the capacity of the existing road link  $l$ :

$$\frac{dC_l^e(t)}{dt} = U_l^c(t), \quad \forall l, \tag{8}$$

where  $U_l^c(t)$  is a continuous control variable  $u(t)$  for *new capacity* with dimension [ $\text{pcu h}^{-1} \text{T}^{-1}$ ], where  $T$  is the time-scale of optimisation. Hence,  $U_l^c(t)$  is changing the functionality of the road link as in a Continuous Network Design Problem (CNDP).

To calculate the total non-point emissions for pollutant  $p \in \mathcal{P}$  the following equation may be used:

$$\frac{dE(t)}{dt} = \gamma_1 \left( \hat{E}(t) - E(t) \right). \tag{9}$$

The disequilibrium level of emissions  $\hat{E}(t)$  is calculated as the number of mode specific trips multiplied with a mode and pollutant-specific emission factor  $\epsilon_{m|p}$ , link length  $d_l$ , and a speed factor (that assumes emissions are lower at higher speeds in network), using parameter  $\beta_{sm|p}$ , which is also applied in Kim and Hoskote [7]:

$$\begin{aligned} \hat{E}(t) = & \alpha_4 \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}^i} \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}_{ij}} \sum_{l_r \in r} \theta_{1m} (Q_i(t) D_{ij}(t) G_{ijmr}(t)) \\ & \epsilon_{m|p} d_{l_r} \left( \frac{d_{l_r}}{\tau_{l_r}(t)} \right)^{-\beta_{sm|p}}, \end{aligned} \tag{10}$$

with route travel time  $\tau_{l_r}(t)$  and route distance  $d_{l_r}$ .

Eqn. (10) is formulated for use of one, possibly dominant, pollutant type only. If necessary, a composite emission factor representing several types of pollutants  $p \in \mathcal{P}$  can be introduced. The environmental capacity is represented through  $E^*$ , and is implemented in the model by use of an exponential penalty function that penalises exceeding the threshold  $E^*$ .

The cost criterion in this example (several others can and have been formulated [1]) aims to keep or bring the level-of-service for a road link or all links in the network at a certain volume-over-capacity level  $\delta_l$  (for example  $\delta_l = 0.80$ , representing high density stable traffic flow) over the time horizon  $(t_0 - t_1)$ , which is, again following [5]:

$$\min \int_{t_0}^{t_1} \sum_l \left( \frac{V_l(t)}{C_l^e(t)} - \delta_l \right)^2 dt \equiv \max - \int_{t_0}^{t_1} \sum_l \left( \frac{V_l(t)}{C_l^e(t)} - \delta_l \right)^2 dt. \tag{11}$$

The implication of this cost criterion is that the control paths are chosen as such that all existing infrastructure is used homogeneously, that is links with a low

volume-over-capacity level  $V_l(t)/C_l^e(t)$  will be increasingly used to reach level  $\delta_l$ , whereas the contrary applies to links with a high volume-over-capacity level.

The solution to the optimal control problem has been characterised using the Pontryagin Maximum Principle [8]. The Maximum Principle is very useful as it provides a set of necessary and sufficient conditions that are required to find the state variable trajectories, in other words transition paths for the state variables, and the optimal set of controls, which maximise the objective function. These conditions are derived from the Lagrangean equation and also provide other important, call it economic, information on marginal valuation of the state variable at some point in time in terms of costs and benefits, through the costate variables.

### 3 Example

The small network in figure 1 is used to demonstrate the dynamic model. Here, an elastic travel demand model is considered, which is sensitive to changes in network performance. Furthermore, the objective function aims at a 80% level-of-service, implying high-density stable traffic flow. An emission state constraint to  $E(t)$  (in this case  $E^* = 81\text{kg CO emissions}$ ) is applied using a penalty function.

Applying this model, one would expect a control path for  $U_l^c(t)$  that minimises the cost criterion as much as possible, while taking consideration of the elastic demand, though not violating the emission constraint. On urban roads, the emission factor is monotonously decreasing with increasing speeds, which can a/o's be accomplished through capacity enhancement.

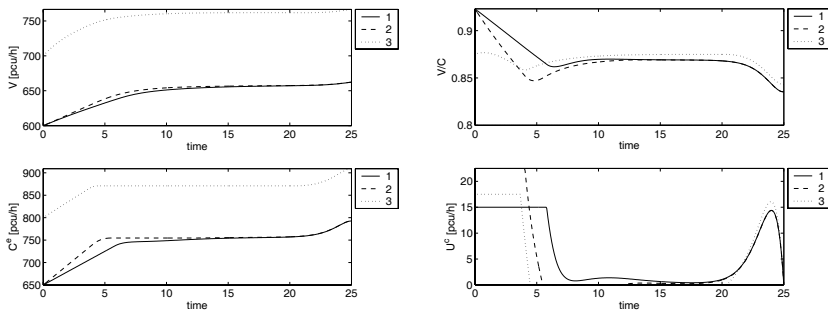


Figure 2: State variables  $V_l(t)$  and  $C_l^e(t)$ , as well as Level-Of-Service  $V_l(t)/C_l^e(t)$  and control variable  $U_l^c(t)$ .

In figure 3 the transition paths for the state variables  $V_l(t)$  and  $C_l^e(t)$  are depicted and show an increase and stabilisation of these variables over time. It can be noticed that the cost criterion for the volume-over-capacity ratios is not minimised at:  $\delta_l = 0.8$ , but slightly above that, which is obviously caused by the emissions constraint, that forces the control path for  $U_l^c(t)$ . These control paths



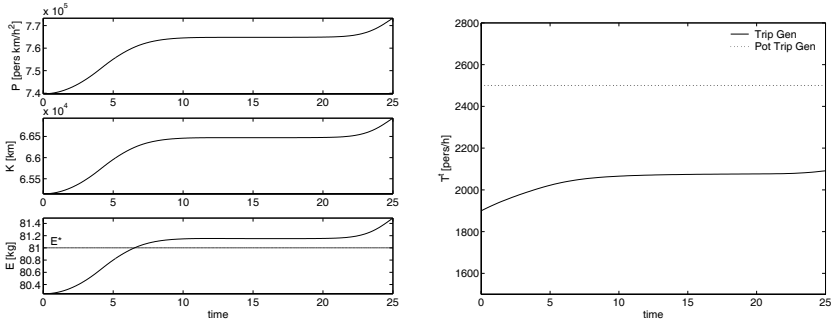


Figure 3: State variables  $P(t)$ ,  $K(t)$  and  $E(t)$ , as well as (potential) trip generation  $T^t$ .

switch between lower and upper bounds, showing the bang-bang nature of the model. A so-called smoothing effect of  $\alpha_7$  that minimises spending on  $U_l^c(t)$  is slightly visible as well. The smoothing parameter is:  $\alpha_7 = 1.0 \times 10^{-6}$ , whereas:  $\chi_1 = 0.05$ .

Figure 3 shows the transition paths for some external effects (person throughput  $P(t)$ , kilometres travelled in the network  $K(t)$ , and emissions  $E(t)$ ). Alike the general character of the penalty function, the target  $E^*$  is slightly exceeded, in particular near the endpoint. Furthermore, the increase in generated trips over time, due to improved accessibility, is shown. The revealed travel demand over time increases to a level close to  $2100 \text{ pers h}^{-1}$ , because of the elasticity function  $\hat{V}_l(t)$ . Also the potential travel demand  $Q(t)$  is shown, which would be revealed if conditions were improved to free-flow conditions and the environmental capacity was set sufficient high.

### 4 Conclusion

It is shown that the notion and definition of sustainable development can be integrated into urban transport planning by advocating a comprehensive decision-making that anticipates and manages scarce resource use, while developing the transport system in terms of quality of access and/or person throughput, hence stimulating social-economic development, based on an explicit and integrated transport policy objective.

This is done by characterising the sustainable transport development problem as a dynamic and constrained optimisation problem, while incorporating a quantified transport policy objective as well as resource bounds.

By applying this to an example network, the use of a dynamic optimisation model that shows where, when and by how much transport measures should be applied in order to steer the system in the direction of the maximised sustainable

transport policy objective, is demonstrated. A more effective and efficient (scarce) resource allocation can be observed. Because of the dynamics in the model there is continuous feedback between the different sub models of the transport model, hence also incorporating generated demand. Implications of measures in terms of transport system performance, but also the number of generated trips in time and total emissions, are thus known. In addition, ‘marginal costs’ associated with erecting boundaries to some control and state variables are explicitly known to the transport policy maker.

## References

- [1] Zuidgeest, M.H.P., *Sustainable Urban Transport Development: A Dynamic Optimisation Approach*. TRAIL Thesis Series T2005/3, The Netherlands TRAIL Research School: The Netherlands, 2005.
- [2] WCED, *Our Common Future*. World Commission on Environment and Development (WCED): Oxford, England, 1987.
- [3] Gudmundsson, H. & Höjer, M., Sustainable development principles and their implications for transport. *Ecological Economics*, **19**, pp. 269–282, 1996.
- [4] Replogle, M., Sustainability: a vital concept for transportation planning and development. *Journal of Advanced Transportation*, **25(1)**, pp. 3–18, 1991.
- [5] Donaghy, K.P. & Schintler, L.A., Managing congestion, pollution, and pavement conditions in a dynamic transportation network model. *Transportation Research*, **3D(2)**, pp. 59–80, 1998.
- [6] De la Barra, T., *Integrated Land Use and Transport: Decision Chains and Hierarchies*. Cambridge University Press: Cambridge, New York, USA, 1989.
- [7] Kim, T.J. & Hoskote, N.G., Estimating mobile source pollutant emission: methodological comparison and planning implications. *Environmental Monitoring and Assessment*, **3**, pp. 1 – 12, 1983.
- [8] Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V. & Mischenko, E.F., *The Mathematical Theory of Optimal Processes*. John Wiley & Sons: New York, USA, 1962.

