USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No. 031OY02

Incorporating Image-Based Traffic Information for AADT Estimation: Operational Developments for Agency Implementation and Theoretical Extensions to Classified AADT Estimation

By

Mark R. McCord
Professor of Civil and Environmental Engineering and Geodetic Science
The Ohio State University
mccord.2@osu.edu

and

Prem Goel, Professor of Statistics
The Ohio State University
goel.1@osu.edu

DISCLAIMER

Funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program, and by the Ohio State University College of Engineering's Transportation Research Endowment Program.

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

NEXTRANS Project No. 031OY02 Final Report, December 18, 2011

Incorporating Image-Based Traffic Information for AADT Estimation: Operational Developments for Agency Implementation and Theoretical Extensions to Classified AADT Estimation

Introduction

Average annual daily traffic (AADT) is perhaps the most fundamental measure of traffic flow. The data used to produce AADT estimates are typically collected by in-highway traffic counters operated by state Department of Transportation (DOT) traffic monitoring crews who must cover thousands of highway segments in their statewide systems on a continual basis. In-highway traffic counters can be dangerous to traffic crews and disruptive to traffic. In addition, the availability of limited resources, coupled with the large number of highway segments spread across the expansive geographic regions of the state, requires that the state DOTs can only collect short-term, sample volumes for the majority of the highway segments. Moreover, not all segments can be sampled every year, and some traffic counts will have been collected several years before the AADT of the segment is estimated.

In a first year project, results of empirical studies demonstrated more accurate AADT estimation when using a proposed method to combine older, traditionally collected traffic count data with traffic information contained in more recently obtained air photos. Software components were also developed to allow many of the calculations to be performed automatically. Additional empirical studies were conducted this year, a refinement to the estimation of an important input value using image-based traffic information was developed, and proof-of-concept software was installed and used at the Ohio Department of Transportation.

Findings

An empirical study using images and ground-based data collected by the Ohio Department of Transportation (ODOT) demonstrated the stability of a previously assumed parameter used when combining image-based information and traditional ground-based traffic counts to estimate AADT. Using twelve ODOT images of highway segments equipped with Automatic Traffic Recorders (ATRs), the standard deviation of the ratio of the estimated AADT produced from image-based information to an estimate of the true AADT produced from ATR data was calculated. The calculated standard deviation was almost identical to the value produced in a previously conducted study using different images.

The refinement of the standard deviation estimate proposed in this study is based on information available in the image. An empirical study showed that AADT estimates produced using “image-based”

estimates of the standard deviation parameter were better than AADT estimates produced when using the default value of this parameter. The improvement was slight, but greater improvements may be exhibited when using images of segments with conditions that differ more substantially from those corresponding to the imaged segments used in this empirical study, which produced an estimated standard deviation value that is very close to the default value.

Extensive help from ODOT personnel was required to allow installation of previously developed software modules on the ODOT computer system. However, this appears to require only a "one-time investment," and should not be a problem for operational use of an image-based approach to AADT estimation if state DOTs committed to implementation of the software. Use of the installed software system highlighted that it would be difficult to use if images were not georeferenced, and the image-based approach to AADT estimation might only be cost effective if a DOT collects georeferenced images on a regular basis. In addition, many images in the ODOT database were actually mosaics of multiple overlapping images taken at different times. If the proposed AADT estimation approach is to be pursued in the future, it would be important to have access to the original images, and not only mosaics of multiple images. Different approaches could be developed and investigated to address correlation when the images are taken within only a few seconds of each other.

Recommendations

The empirical findings continue to indicate that AADT estimates can be improved by incorporating image-based information using the methodology proposed in this research. The installation and use of proof-of-concept software at the Ohio Department of Transportation indicates the feasibility of developing a software system for operational use if georeferenceced images are collected by state DOTs and if original images, and not only mosaics of images, are accessible. To motivate further progress toward implementation, additional trial use at multiple state Departments of Transportation is recommended. Such studies would identify operational issues for sustained and efficient use. Further empirical studies of the performance of the proposed method for refining the standard-deviation parameter estimate using image-based traffic information are also recommended.

Contacts

For more information:

Mark R. McCord
Principal Investigator
Deptt of Civil & Environ. Engineering and Geodetic Science
The Ohio State University
mccord.2@osu.edu

Prem Goel
Co-Principal Investigator
Department of Statistics
The Ohio State University
goel.1@osu.edu

NEXTRANS Center
Purdue University - Discovery Park
2700 Kent B-100
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9729
(765) 807-3123 Fax

www.purdue.edu/dp/nextrans

This page intentionally left blank.

NEXTRANS Project No. 031OY02

Incorporating Image-Based Traffic Information for AADT Estimation: Operational Developments for Agency Implementation and Theoretical Extensions to Classified AADT Estimation

by

Mark R. McCord
Professor of Civil and Environmental Engineering and Geodetic Science
The Ohio State University
mccord.2@osu.edu

and

Prem Goel, Professor of Statistics
The Ohio State University
goel.1@osu.edu

Report Submission Date: December 18, 2011

ACKNOWLEDGMENTS AND DISCLAIMER

The investigators gratefully acknowledge the assistance of Yufang Zhang, who, as a graduate research associate at OSU, was instrumental in conducting the reported empirical studies and installing the software at ODOT during many onsite visits. In addition, the installation and demonstration of the software would not have been possible without the additional assistance of the Ohio Department of Transportation, especially, Dave Gardner (Traffic Monitoring - Technical Services), Dave Blackstone (GIS - Technical Services), John Ray (Aerial Engineering), Chuck Bernthold (Aerial Engineering), Garret Staat (Technical Services), and Rob Watson (Technical Services).

Funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program, and by the Ohio State University's College of Engineering Transportation Research Endowment Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented, and not of any collaborator or funding institution. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Average annual daily traffic (AADT) is perhaps the most fundamental measure of traffic flow. The data used to produce AADT estimates are typically collected by in-highway traffic counters operated by state Department of Transportation (DOT) traffic monitoring crews who must cover thousands of highway segments in their statewide systems on a continual basis. In-highway traffic counters can be dangerous to traffic crews and disruptive to traffic. In addition, the availability of limited resources, coupled with the large number of highway segments spread across the expansive geographic regions of the state, requires that the state DOTs can only collect short-term, sample volumes for the majority of the highway segments. Moreover, not all segments can be sampled every year, and some traffic counts will have been collected several years before the AADT of the segment is estimated.

We previously developed a method to combine the older, traditionally collected traffic data with traffic information contained in more recently obtained air photos in a statistically supported manner designed to produce more accurate estimates of AADT. The appeal of this result is that state DOTs image highways for purposes unrelated to traffic flow analysis and can also easily obtain images of specific highway segments when flying to or from a data collection mission scheduled for other purposes. As such, the marginal cost of obtaining the image information is very low. If a method of combining image-based traffic information with traditionally collected traffic data to improve AADT estimation is implemented, data collection procedures could be adjusted so that the number of costly and dangerous traffic counts is reduced while improving accuracy in estimating AADT. To take advantage of this promising method in practice, it is necessary to demonstrate its potential for better estimation accuracy and to develop an efficient way to use the method on a widespread, repeated basis in an operational setting.

In a first year project, we conducted empirical studies that demonstrated the advantage of the proposed method, developed software components to conduct many of the calculations automatically, and motivated traffic monitoring personnel at the Ohio Department of Transportation (ODOT) to allow us to develop software at ODOT as step toward developing an operational system. In the second year effort reported here, we continued to investigate the quality of AADT estimation using imagery by conducting additional empirical studies, developed a refinement of the estimate of an important input value based on information available in the image, and installed proof-of-concept software at the Ohio Department of Transportation.

PROBLEM

The overall problem to be addressed is that of working toward the development, implementation, and use of a system and a process in which aerial imagery, primarily collected for non-traffic monitoring purposes by state DOTs, is used to improve AADT estimates.

APPROACH

Our efforts in the year 2 effort were devoted to two major thrusts – empirical investigations and software implementation.

Thrust 1 - Empirical investigations of AADT estimation performance: This thrust consisted of two components: (1) an investigation of the assumed value of an important input variable used in our AADT estimation method; (2) an investigation of AADT estimation performance when refining this input value using information available in the imagery.

Thrust 2 - Implementation of software: In this thrust we worked with personnel at the Ohio Department of Transportation (ODOT) to implement software on the ODOT system and use it to estimate AADT from an ODOT image in a manner that would emulate operational use.

METHODOLOGY

Incorporating Image-based Information in AADT Estimation

Details of the traditional approach to estimating AADT and the approach we have developed that combines information in air photos with traditionally collected ground counts to provide an improved AADT estimate can be found in McCord and Goel (2009). To summarize, in the traditional approach AADT in year *y* on highway segment *s* is estimated as the average of a set of 24-hour traffic volumes produced from "coverage counts" (traffic counts that are scheduled in a DOT traffic data collection program so as to "cover" the highway network on a multiyear basis) that are "deseaonalized" by factors that account for the temporal variability in traffic attributable to the day-of-week and the month-of-year on which the coverage counts were taken. We denote this traditional estimate as $AADT^{C}_{s}(y)$, where the superscript *C* represents that the estimate is produced from coverage counts.

Because of limitations in the supply of equipment and personnel, coverage counts cannot be obtained on all segments every year. To estimate AADT in some year *y'* after the year *y* in which the coverage counts were obtained, it is typical to multiply $AADT^{C}_{s}(y)$ by a growth factor $GF_{s}(y, y')$ that accounts for the estimated growth in traffic between year *y* and year *y'*. This growth factor is generally estimated from traffic data produced by Automatic Traffic Recorders (ATRs) on a set of segments where traffic patterns are believed to be similar to those of the segment *s* for which the AADT is being estimated. (ATRs are permanently installed traffic recorders designed to collect traffic 24 hours per day, 365 days per year on a small subset of highway segment.) We denote this resulting estimate as $AADT^{CG}_{s}(y', y)$, where the superscript *CG* indicates that both coverage counts and a growth factor are used in the estimation.

The approach we have developed to integrate the more contemporary information available in an image of segment *s* taken in hour-of-the-day *h* on day-of-the-week *d* in month-of-the-year *m* in year *y'* first produces an "image-based" AADT calculated as:

$$AADT^{I}_{s}(\gamma'; h,d,m) = N^{veh}/L \times U \times 24 \times F^{H}_{s}(h,d) \times F^{MD}_{s}(m,d) \quad (1)$$

where N^{veh} is the number of vehicles in the image on segment *s, L* is the length of the segment considered in the image, *U* is the average speed on the segment (which could depend on the traffic density of the segment), $F^{h}_{s}(h, d)$ is a factor used to convert an hourly traffic volume occurring during hour *h* and day-of-week *d* on segment *s* to an average hourly volume for the day, and $F^{MD}_{s}(m,d)$ is the seasonal factor used to convert a 24-hour traffic volume occurring on day-of-week *d* and month-of-year *m* to an estimate of an average volume for the year.

In the proposed approach, the two estimates, $AADT^{CG}_{s}(\gamma', \gamma)$ and $AADT^{I}_{s}(\gamma)$, of AADT in year *y'* are combined using a weighted average to produce the proposed improved estimate of AADT in year *y'*:

$$AADT^{CGI}_{s}(\gamma', \gamma) = w \times AADT^{I}_{x}(\gamma') + (1-w) \times AADT^{CG}_{s}(\gamma', \gamma). \quad (2)$$

The superscript *CGI* indicates that the estimate incorporates the coverage counts (*C*), the growth factor (*G*), and the image (*I*). The weight *w* used in Equation (2) is derived from:

$$w = [(\sigma^{C})^2 + (\sigma^{G})^2] / [(\sigma^{C})^2 + (\sigma^{G})^2 + (\sigma^{I})^2] \quad (3)$$

where $(\sigma^{C})^2$ is the variance of the ratio the coverage count-based AADT estimate in year *γ* (the year in which the coverage counts were obtained) to the true AADT in year *γ*, $(\sigma^{G})^2$ is the variance of the ratio of the estimated growth factor for segment *s* between years *γ* and *γ'* to the true growth factor for the segment, and $(\sigma^{I})^2$ is the variance of the ratio of the image-based AADT estimate in year *γ'* (the year in which the image was obtained) to the true AADT in year *γ'*. As explained in McCord and Goel (2009) and Jiang *et al.* (2006), $(\sigma^{C})^2$ and $(\sigma^{G})^2$ can be estimated from available ATR data, and a default value of $(\sigma^{I})^2$ determined from empirical studies can be used in Equation (3) to determine the value of *w*.

Further Empirical Investigations of Estimation Performance

In Jiang *et al*., (2006), the $AADT^{CGI}_{s}(\gamma', \gamma)$ estimate is argued to be conceptually more accurate than the traditional estimate $AADT^{CG}_{s}(\gamma', \gamma)$ and shown to perform better in a simulation study. In our first year study (McCord and Goel, 2009), we used 12 images of 6 highway segments collected by the Ohio Department of Transportation (ODOT) Aerial Engineering section, the corresponding traffic counts obtained by the ODOT Traffic Monitoring section, and the prototype software we developed to demonstrate the improved performance of $AADT^{CGI}_{s}(\gamma', \gamma)$ in an empirical study.

In both the simulation and empirical studies, a default value of σ^{I} (the standard deviation corresponding to the variance $(\sigma^{I})^2$) was used in Equation (3) to determine the weight of *w* of the image-based estimate (relative to the traditional AADT estimate) that is used in Equation (2) to produce $AADT^{CGI}$. The default

value of σ^I was based on empirical comparisons between $AADT^I$ produced from Equation (1) and corresponding AADT values either published by the state DOT or determined from ATR data (McCord, *et al.,* 2003). Using 22 comparisons, an estimated value of 0.17 was produced. In the study reported here, we used the twelve ODOT images, none of which was used in the McCord, *et al.* (2003) study, and corresponding traffic monitoring data to investigate this default estimate of σ^I. In addition, we conducted an empirical study in which we used a refined estimate of σ^I based on information that can be obtained in the image. We explain the methodology of each of these studies next.

Default σ^I Value

We used the same twelve images used in our year 1 study to investigate the reliability of the previously proposed default value of σ^I. All the images were obtained by ODOT Aerial Engineering in 2005 and each image contained one highway segment on which an ODOT ATR was located. We used Equation (1) to calculate $AADT^I$ in 2005 for each of the ATR-equipped segments in the images and used the 2005 ODOT ATR data with the AASHTO method (AASHTO, 1992) to produce estimates of the true AADTs in 2005 for the segments. (We call these "estimates" of the true AADT, since there are errors or missing data in the ATR dataset.) We then formed the ratio of each of the twelve estimated $AADT^I$s (one for each image-ATR equipped segment pair) to the true AADT for the segment and calculated the standard deviation of the set of ratios. We note that the images consisted of sets of two images on six different roadway segments equipped with ATRs. The two images for a given segment were taken only a few seconds apart and contained many of the same vehicles. As such, the twelve ratios of $AADT^I$ to the corresponding true AADT cannot be considered independent. Nevertheless, the standard deviation would still be meaningful, especially given the sometimes very different values of $AADT^I$ that were obtained from the two images of the same segment.

AADT Estimation Using Image-based Estimate of σ^I

Image-based estimate of σ^I: The underlying concept in estimating the AADT for a segment from an image of the segment, as given in Equation (1), is to first obtain the traffic density on the segment from the image and determine an average hourly flow rate from the density and an assumed average speed using the fundamental relation of macroscopic traffic flow (see, *e.g.,* Mannering *et al.*, 2009), namely, that flow rate equals density times speed. The flow rate, which can be used to produce an estimate of the hourly volume, is converted to an estimate of the 24-hour volume for the day by multiplying by 24 hours and the hourly factor. The estimate of the 24-hour volume is then converted to an estimate of AADT by using the traditional seasonal factors that account for the day-of-week and month-of-year when the traffic data were imaged. In this way, the error in estimate $AADT^I$ can be considered to be comprised of an error in converting the observed density to an estimate of the hourly volume, an error in converting the hourly volume to an estimate of the 24-hour volume, and in error in converting the 24-hour volume to an estimate of the annual average daily volume.

According to the arguments in the Appendix, the variance $(\sigma^I)^2$ of the ratio of $AADT^I$ to the true AADT used in Equation (3), can be approximated as

$$(\sigma^{I})^2 \approx \sigma_D^2+\sigma_H^2+\sigma_T^2, \quad (4)$$

where σ_D^2 is the variance of the ratio of the estimated AADT, conditional on knowing the true 24-hour volume for a day, to the true AADT; σ_H^2 is the variance of the ratio of the estimated 24-hour volume for a given day, conditional on knowing the true volume in a given hour on the day, to the true 24-hour volume for the day; and σ_T^2 is the variance of the ratio of the estimated volume in a specified hour, conditional on knowing the true volume in a sub-interval of the hour of specified duration, to the true volume in the given hour.

Estimates of σ_D^2 and σ_H^2 can be developed by determining sample variances of the corresponding ratios using data from ATRs believed to behave similarly, in terms of temporal variability, to the segment *s* for which the AADT is to be estimated. More specifically, to estimate σ_D^2, a daily (24-hour) ATR volume would be converted to an estimate of an annual average by using seasonal factors, as is traditionally done. The estimate of the annual average would then be divided by the "true" AADT (determined from the ATR data) to form an empirical ratio. Repeating for all relevant days (*e.g.,* all days, or all weekdays, or all Tuesdays, Wednesdays, and Thursdays) would form a set of empirical ratios. The variance of this set of empirical ratios can be used to approximate σ_D^2.

Similarly, to estimate σ_H^2, an hourly volume available from the ATR data can be converted to a 24-hour volume estimate by multiplying by 24 and using the hourly factor corresponding to the hour. The 24-hour volume on that day is also available from the ATR data. (The 24-hour volume is the sum of the twenty-four hourly volumes on that day.) The ratio of the estimated 24-hour volume to the true 24-hour volume can then be formed. Repeating for different hours, different days, and different hours would form a set of ratios, and the sample variance of the set of ratio can be used to approximate σ_H^2.

To estimate σ_T^2 we consider that the vehicles in the image can be converted to a sub-hourly volume or traffic count of duration t^{dur}. The "equivalent count duration," which is typically on the order of a minute to a few minutes (*i.e.*, a few hundredths of an hour) can be obtained from the assumed average vehicle speed and the length of the segment in the image (see, Jiang *et al.*, 2006; McCord *et al.*, 2003). Conditional on knowing the hourly volume V^H, we assume that the number of vehicles that would "appear" in the subhourly period of duration t^{dur} is binomially distributed, with the hourly volume V^H and the proportion of the hour t^{dur} (= t^{dur} [hrs] / 1 [hr], assuming t^{dur} is in units of hours) covered by t^{dur} representing the number of "trials" and the "probability of success," respectively. In the Appendix, we show that the variance σ_T^2 of the ratio of the hourly volume – estimated from the number of vehicles N^{veh} in the image, the length *L* of the segment in the image, and the average speed *U* of the vehicles on the segment – to the true hourly volume can be approximated as:

$$\sigma_T^{\ 2} = (1-L/U)/N^{veh}. \quad (5)$$

As would be expected, Equation (5) shows that the variance $\sigma_T^{\ 2}$ of the estimated hourly volume is reduced as the length *L* of the segment in the image increases (which leads to a longer "equivalent count interval duration," everything else being equal) and as the number of vehicles N^{veh} in the image increases (which is consistent with a larger number of observations, everything else being equal).

In summary, the number of vehicles in the image and the length of the segment in the image can be used to determine an estimate of σ_T^2, and ATR data can be used to determine estimates of σ_D^2 and σ_H^2. These estimates can be summed (see Equation (4)) to produce an "image-based" estimate of $(\sigma^I)^2$. This image-based estimate can then be combined according to Equation (3) with the estimates of $(\sigma^C)^2$ and $(\sigma^G)^2$ to produce a value of *w* that weights the imaged-based and traditional AADT estimates according to Equation (2).

Design of empirical study: We conducted an empirical study of the ability of the image-based estimate of $(\sigma^I)^2$ to improve the AADT estimate $AADT^{CGI}$, compared to using a default value of $(\sigma^I)^2$. In our empirical study, we used the same twelve images used in our year 1 study, estimated $AADT^{CGI}$ on the segment *s* in the image upon which an ATR was located using a weight *w* in Equation (2) derived from the image-based value of $(\sigma^I)^2$ and the default value of $(\sigma^I)^2$, and compared the results to the estimate of the true AADT on segment *s* obtained from the ATR data. As mentioned above, the twelve images, all of which were obtained in 2005, were comprised of pairs of two images taken a few seconds apart of six different ATR-equipped segments. In this way, there were six segments investigated in the study, with two images for each segment.

As in the first year study (McCord, *et al.,* 2009), we used ATR data on segment *s* to determine $AADT^{CGI}$ values based on coverage counts (generated from ATR data) on segment *s* from 2003 and 2004. The same general approach used in our first year study was used to produce the true 2005 AADT values, which we denote $AADT^{true}$, and the $AADT^{CGI}$ values using the default $\sigma^I = 0.20$ (rounded from 0.17) based on generated coverage counts from 2003 and 2004. We denote the estimate based on this default value $AADT^{CGI(def)}$. The value of $AADT^{CGI}$ determined using the image-based estimate of σ^I, which we denote $AADT^{CGI(img)}$, differed from $AADT^{CGI(def)}$ only in that the value of σ^I was determined from equation (5) using information from the image, rather than a default value of 0.20. We note that estimated AADT values produced in this report differ slightly from those produced in the year one report. Different individuals produced the estimates in the two years. Image-based AADT would depend on the length *L* of the segment in the image, and different individuals would delimit this length differently. Furthermore, the different individuals would have processed the ATR data slightly differently when determining which data to include and which to exclude when producing the various AADT estimates – estimates of the "true" AADT values, growth factors, σ^C, and σ^G. In addition, for one segment, we considered two directional AADT this year, but considered only one directional AADT in the year one project. However, when we repeated the year-one comparisons with the values produced this year, we again saw that the AADT estimates $AADT^{CGI}$ that incorporated image information outperformed the estimates that did not incorporate image information. (Specifically, $AADT^{CGI}$ outperformed $AADT^{CG}$ and a second estimate called $AADT^{c}$ that was investigated in the year-one study.)

Implementation of Estimation Software at Ohio Department of Transportation

We worked with section leaders in Traffic Monitoring and Geographic Information Systems (GIS) sections at the Ohio Department of Transportation (ODOT) to connect the software modules developed

in the year one effort and install the connected modules on the ODOT computer system. We then applied the software to an air photo obtained from the ODOT Aerial Engineering section.

FINDINGS

Empirical Study of Default σ^I Value

In Table 1, we list the image number, the ODOT number for the ATR-equipped segment in the image, the functional class (FC) of the segment, the year in which the image was obtained (which was 2005 for all segments, and the same year for which the AADT was calculated), the length *L* of the image in miles that appeared in the image, the number of vehicles N^{veh} in the image, the assumed vehicle speed *U* at the time the image was taken, the image-based AADT estimate $AADT^I$ produced from Equation (1), the estimate $AADT^{true}$ of the true AADT produced from the ATR data, and the ratio of the image-based to true AADT estimates. As seen at the bottom of the table, the mean of the $AADT^I$ values is approximately 1 (0.9675). Of interest for this study, the standard deviation is 0.17518. This sample standard deviation is amazingly close to the value of 0.17 produced in McCord, *et al.* (2003) and which formed the basis for the default value used in previous empirical studies. We note again that none of the twelve images used in this study had been used in the McCord, *et al.* (2003) empirical study.

Table 1. Ratio of Image-based Estimate $AADT^I$ to True AADT for Twelve Images of Empirical Study

Image #	ATR #	FC	Image Yr	L [mi]	N^{img}	U [mph]	$AADT^I$	$AADT^{true}$	$AADT^I/AADT^{true}$
11273-1-1	767	12	2005	0.698	11	70	20615	31926	0.646
11273-1-2	767	12	2005	0.713	18	70	33004	31926	1.034
11273-2-3	707	1	2005	0.655	15	75	32896	30424	1.081
11273-2-4	707	1	2005	0.657	15	75	32781	30424	1.077
11273-4-5	601	11	2005	0.686	34	70	65090	94867	0.686
11273-4-6	601	11	2005	0.705	59	70	109850	94867	1.158
11273-5-7	752	11	2005	0.685	49	70	93903	94301	0.996
11273-5-8	752	11	2005	0.707	46	70	85377	94301	0.905
11273-3-9	121	11	2005	0.655	64	70	128170	123992	1.034
11273-3-10	121	11	2005	0.663	64	70	126810	123992	1.023
11273-6-11	140	11	2005	0.683	36	70	69235	58712	1.179
11273-6-12	140	11	2005	0.681	24	70	46263	58712	0.788
								Average	0.967
								St Dev	0.175

The closeness of the σ^I estimate produced in this empirical study to the value produced in a previous study using a completely different set of images, segments, and years of AADT estimation is striking and supports the promise of using this estimation approach with the use of a default value.

Empirical Results of AADT Estimation Using Image-based Estimate of σ^J

To compare the performance of the AADT estimate $AADT^{CGI(img)}$ using an image-derived estimate of σ^J to the performance of the AADT estimate $AADT^{CGI(def)}$ using the default value $\sigma^J = 0.20$, we considered the three measures of performance used in our first year study. Specifically, we formed

- the mean absolute relative error *MARE* between the $AADT^{CGI(img)}$ estimates and the true AADT and between the $AADT^{CGI(def)}$ estimates and the true AADT. (Lower MARE is better.)
- the proportions *Prop(ARE<0.10)* of $AADT^{CGI(img)}$ estimates and of $AADT^{CGI(def)}$ estimates within 10% (a commonly used target) of the true AADT. (Higher proportion is better.)
- the proportion of times $Prop(ARE^{CGI(img)} < ARE^{CGI(img)})$ that $AADT^{CGI(img)}$ was closer than $AADT^{CGI(def)}$ to the true AADT. (A proportion greater than 0.5 indicates that $AADT^{CGI(img)}$ outperformed $AADT^{CGI(def)}$ more often than $AADT^{CGI(def)}$ outperformed $AADT^{CGI(img)}$.)

As in our first year study, we produced estimates of the 2005 AADT using 24-hour volumes obtained from ATRs in 2004 and in 2003 to emulate coverage count data obtained in those years.

The results are presented in Table 2. In addition to presenting the *MARE*, *Prop(ARE<0.10),* and $Prop(ARE^{CGI(img)} < ARE^{CGI(img)})$ values, we also indicate which estimate performed better according to the indicated measure for the specific image. An entry of zero for each estimate indicates that the two estimates performed identically on the measure. The sums, across the images and the years, of the number of times each estimate outperformed the other are presented at the bottom of the table.

The results show that $AADT^{CGI(img)}$ outperformed $AADT^{CGI(def)}$ more often than the contrary for all three measures of performance. The sample size was small (indeed, it is difficult to obtain the data that allow the type of empirical study we have designed and conducted), and $AADT^{CGI(img)}$ outperformed $AADT^{CGI(def)}$ only slightly more often than the contrary. Therefore, additional studies would be needed before concluding that Equation (5) should be used to incorporate image-based information in the estimation of σ^J. Nevertheless, the empirical study of the default value of σ^J presented above showed that the default value was a very good estimate for this set of images. It is possible that the image-based estimate of σ^J would show a greater improvement for a more diverse set of images. Therefore, the results are supportive of the potential of using information in the image, either through Equation (5) or by some other means, to improve the weight placed on the image-based estimate of AADT relative to the traditional estimate when producing a combined AADT estimate

Table 2. Summary Measures of Performance for Image-based AADT Estimates using Image-based Information to Estimate σ^I ($AADT^{CGI(img)}$) and Default Value of σ^I ($AADT^{CGI(def)}$), by Image

	Image name	ATR#	MARE		Better MARE		Prop(ARE < 0.10)		Better (ARE < 0.10)		Prop($ARE^{CGI(img)}$ <$ARE^{CGI(def)}$)	Better Estimate	
			$AADT^{CGI(img)}$	$AADT^{CGI(def)}$	$AADT^{CGI(img)}$	$AADT^{CGI(def)}$	$AADT^{CGI(img)}$	$AADT^{CGI(def)}$	$AADT^{CGI(img)}$	$AADT^{CGI(def)}$		$AADT^{CGI(img)}$	$AADT^{CGI(def)}$
Coverage Counts from 2004	11273-1-1	767	0.0297	0.0355	1	0	0.970	0.970	0	0	0.720	1	0
	11273-1-2	767	0.0286	0.0277	0	1	0.980	0.980	0	0	0.450	0	1
	11273-2-3	707	0.0632	0.0629	0	1	0.854	0.854	0	0	0.521	1	0
	11273-2-4	707	0.0631	0.0626	0	1	0.854	0.854	0	0	0.469	0	1
	11273-4-5	601	0.0523	0.0489	0	1	0.907	0.907	0	0	0.070	0	1
	11273-4-6	601	0.0863	0.0848	0	1	0.698	0.698	0	0	0.035	0	1
	11273-5-7	752	0.0540	0.0542	1	0	0.891	0.891	0	0	0.970	1	0
	11273-5-8	752	0.0481	0.0482	1	0	0.911	0.911	0	0	0.901	1	0
	11273-3-9	121	0.0287	0.0289	1	0	0.979	0.979	0	0	0.474	0	1
	11273-3-10	121	0.0283	0.0286	1	0	0.989	0.979	1	0	0.611	1	0
	11273-6-11	140	0.0409	0.0399	0	1	0.951	0.961	0	1	0.363	0	1
	11273-6-12	140	0.0537	0.0562	1	0	0.902	0.902	0	0	0.824	1	0
Coverage Counts from 2003	11273-1-1	767	0.0459	0.0567	1	0	0.931	0.908	1	0	0.839	1	0
	11273-1-2	767	0.0385	0.0363	0	1	0.943	0.943	0	0	0.241	0	1
	11273-2-3	707	0.0518	0.0518	1	0	0.951	0.951	0	0	0.646	1	0
	11273-2-4	707	0.0517	0.0516	0	1	0.951	0.951	0	0	0.598	1	0
	11273-4-5	601	0.0943	0.0892	0	1	0.692	0.769	0	1	0.000	0	1
	11273-4-6	601	0.1359	0.1346	0	1	0.077	0.077	0	0	0.000	0	1
	11273-5-7	752	0.0529	0.0530	1	0	0.907	0.907	0	0	0.860	1	0
	11273-5-8	752	0.0457	0.0457	1	0	0.907	0.907	0	0	0.977	1	0
	11273-3-9	121	0.0427	0.0436	1	0	1.000	1.000	0	0	0.800	1	0
	11273-3-10	121	0.0431	0.0439	1	0	1.000	1.000	0	0	0.800	1	0
	11273-6-11	140	0.0545	0.0527	0	1	0.939	0.939	0	0	0.204	0	1
	11273-6-12	140	0.0741	0.0768	1	0	0.684	0.653	1	0	0.888	1	0
				Sum	13	11		Sum	3	2	Sum	14	10

Implementation of Software at Ohio Department of Transportation

We successfully installed our estimation modules (McCord, et al, 2009) on ODOT's computer system and estimated AADT from an existing air photo with limited manual intervention. The photo imaged a portion of Licking County State Road 16. Specifically, we georegistered the scanned photo in ODOT's Geographic Information System (GIS) database. We then used a mouse to select the section of roadway over which vehicles would be "counted" and to digitize the location (lat-lon) of each vehicle on the segment. After completing the registration and digitization, the length *L*, segment id, and functional class were automatically available from the GIS data. The most recent AADT estimate based on a coverage count on the segment was obtained from GIS after inputting the date when the image was taken. Presently, date and time information are contained a in a scanned photo of the flight log and input manually, since ODOT has no need to include this information in the image's meta data. If this AADT estimation approach is pursued in the future, it would be straightforward for a state DOT to create a meta data spreadsheet of images with such data included, which would allow this step to be automated. We manually input an average speed *U* of 70 mph. We had previously used ODOT Traffic Monitoring database information to create tables of growth factors, hourly and seasonal factors (F^H and F^{MD}) and of coverage count, growth factor, hourly, and daily standard deviations (σ^C, σ^G, σ_H, and σ_D). The appropriate values were automatically retrieved based on the functional class, date, and time

information associated with the image. The "image-based" information L and N^{veh} were automatically used to determine the sub-hourly standard deviation σ_T. The software then automatically produced the AADT estimate for the year of the image.

In the implementation process, we needed extensive help from ODOT personnel to allow compatibility of our software modules with the ODOT computer system. We believe that this would be a "one-time issue" that would not cause problems for operational use of an image-based approach to improving AADT estimation if state DOTs committed to implementation of the software. We also realized that the system would be difficult to implement if images were not georeferenced. Therefore, we believe that the image-based approach to AADT estimation might only be cost-effective if the DOT collects georeferenced images on a regular basis. We believe that DOTs will be increasingly collecting georeferenced imagery in the future. In addition, we discovered that many of images in the ODOT database were actually mosaics of multiple overlapping images taken at different times. In our study, we chose a road segment that did not cross any overlapping image lines to ensure that the date and time information corresponded to when the vehicles were imaged. If this approach is to be pursued in the future, it would be important to have access to the original images when using this approach. Different approaches could be developed and investigated to address correlation when the images are taken within only a few seconds of each other.

CONCLUSIONS

The results obtained this year build upon those produced in our first year study in supporting the potential of improving AADT estimates by using air photos that either exist in state DOT databases or can be obtained by state DOTs at low cost. The first year study empirically showed the advantage of incorporating an image-based estimate of AADT with an estimate produced from traditional traffic counts in a way that can implemented with no additional data collection. The estimation approach assumed a default value of an input parameter that was derived from a previously conducted empirical study. This year, we conducted an additional empirical study that produced an almost identical value of this input parameter, providing additional confidence in the default value. We also conducted an empirical study indicating that the AADT estimate might be improved by refining the default value with information in the image that is already extracted to produce the image-based AADT estimate.

In the first year study, we also developed software modules to indicate the potential of implementing the approach for operational use. This year, we integrated the modules and installed proof-of-concept software at the Ohio Department of Transportation to investigate the feasibility of efficiently producing an AADT estimate using an ODOT image and ODOT Traffic Monitoring data. Based on this exercise, we believe that an operational system that improves AADT estimates by using existing or easy-to-obtain images could be implemented in state DOTs across the country if the DOTs regularly collect georeferenced images and maintain the original imagery, and not only mosaics of sets of images.

REFERENCES

AASHTO (1992), *AASHTO Guidelines for Traffic Data Programs,* AASHTO, Washington, D.C.

Jiang, Z. McCord, M., and Goel, P. (2006), "Improved AADT Estimation by Combining Information in Image- and Ground-Based Traffic Data," *Journal of Transportation Engineering*, Vol. 132, Issue 7, pp. 523-530.

Mannering, F. Kilareski, S., and Washburn, W. (2009). *Principles of Highway Engineering and Traffic Analysis,* Fourth Edition, John Wiley & Sons, Inc.

McCord,M., and Goel, P. (2009), "Estimating AADT from combined air photos and ground based data: System design, prototyping, and testing," Region 5 U.S. DOT Region V University Transportaiton Center, Nextrans Project No. 002OY01, Final Report.

McCord, M., Yang, Y., Jiang, Z., Coifman, B., and Goel, P. (2003). "Estimating AADT from Satellite Imagery and Air Photos: Empirical Results." *Transportation Research Record*, *1855*, Transportation Research Board, Washington, D.C., pp. 136-142.

APPENDIX - DERIVATION OF σ^I

We call *Noise(img)* the ratio of the image-based AADT estimate $AADT^{img}$ to the true AADT value *AADT:*

$$Noise(img) = AADT^{img} / AADT \quad \text{(A.1)}$$

Rearranging Equation (A.1) to write $AADT^{img} = AADT \times Noise(img)$, we see that *Noise(img)* can be considered a multiplicative error term in estimating *AADT* from a single image.

When producing the image-based AADT estimate $AADT^{img}$, we first obtain the number of vehicles N^{veh} appearing in the imaged portion of the segment. This number can be considered to represent an "equivalent traffic volume" that would be observed during a very short "equivalent count interval" t^{dur} hours. The interval t^{dur} ($0< t^{dur}<1$) is determined by dividing the length of the segment imaged *L* by average speed of vehicles *U* on the segment. The short-term traffic volume N^{veh} is then expanded to an hourly volume estimate, and the estimated hourly volume is expanded to a daily volume estimate by using hourly factors available from ATR data. Finally, the estimated daily volume is deseasonalized to produce $AADT^{img}$ by using the day-of-the-week and month-of-the-year factors currently used in traditional ground-based AADT estimation. These factors can be derived from ATR data. In short, the approach can be represented as:

$$AADT^{img} = (N^{veh} / t^{dur}) \times 24 \times F^{H}(h; d) \times F^{MD}(m,d), \quad \text{(A.2a)}$$

with,

$$t^{dur} = L/U, \quad \text{(A.2b)}$$

where $F^{H}(h; d)$ is a factor used to convert an hourly traffic volume occurring during hour h and day-of-week d on the segment to an average hourly volume for the day, $F^{MD}(m,d)$ is a seasonal factor used to convert a 24-hour traffic volume occurring on day-of-week d and month-of-year m to an estimate of the average volume for the year, and N^{veh}, t^{dur}, L, and U were defined above. We note that substituting (A.2b) in $N^{veh} x (1/ t^{dur})$, which produces an hourly volume estimate, yields the fundamental relation of macroscopic flow stating that (hourly) volume or flow is equal to density (N^{veh} /L) times space mean speed U.

Equation (A.2) shows that there are three main steps involved with converting a single image to an AADT estimate:

(i) converting the "equivalent" short-term traffic count N^{iveh} to an estimate of the hourly volume by dividing the equivalent count interval t^{dur} (in units of hours)

(ii) converting this estimated hourly volume to an estimate of the daily volume by multiplying by 24 and the appropriate hourly factor $F^{MD}(m,d)$

(iii) converting this estimated daily volume to an estimate of AADT by multiplying by the corresponding seasonal factor $F^{MD}(m,d)$

We can think of one noise component (*i.e.*, error) imposed upon the estimate at each step. Specifically, consider three noise components *Noise*(*T*), *Noise*(*H*), and *Noise*(*D*), defined as:

$$Noise(T) = (N^{veh} / t^{dur}) / V^{H}(h,\delta), \quad \text{(A.3a)}$$

$$Noise(H) = V^{H}(h,\delta) \times (24 \times F^{H}(h,d)) / V^{D}(\delta), \quad \text{(A.3b)}$$

$$Noise(D) = V^{D}(\delta) \times F^{MD}(m,d) / AADT, \quad \text{(A.3c)}$$

where $V^{H}(h,\delta)$ is the true hourly volume in hour-of-the-day h on day-of-the-year δ in which the segment was imaged, $V^{D}(\delta)$ is the true daily volume that occurred on day-of-the-year δ, and *AADT* is the true AADT for the year in which the segment was imaged. Similar to *Noise*(*img*) defined in Equation (A.1), all three noise components are the ratios of an estimated volume to the true volume. *Noise*(*D*) is the ratio of the estimated AADT to the true AADT on the segment, where the estimated AADT is developed from the true daily volume by deseasonalizing this daily volume with seasonal factors. In other words, *Noise*(*D*) represents the "random" variation in the true daily volume from the AADT after the daily volume is adjusted by seasonal factors. *Noise*(*H*) is the ratio of the estimated daily volume to the true daily volume, where the estimated daily volume is developed from the true hourly volume expanded to account for all 24 hours in the day and the hourly variability as represented by the hourly factor F^{H}. *Noise(D)* and *Noise(H)* include both the temporal variability that would remain if segment-specific adjustment factors F^{MD} and F^{H} could be obtained and the "spatial variability" that results from the need to estimate a segment's adjustment factors for segments not equipped with ATRs from data collected on supposedly "homogeneous" segments equipped with ATRs. *Noise*(*T*) is the ratio of the estimated hourly volume to the true hourly volume, where the estimated hourly volume is developed by linearly expanding the "observed" t^{dur} count N^{veh}.

With Equations (A.3a) - (A.3c), Equation (A.2) can be rewritten as $AADT^{img} = AADT \times Noise(D) \times Noise(H) \times Noise(T)$ and Equation (A.1) as

$$Noise(img) = Noise(D) \times Noise(H) \times Noise(T) \quad (A.4)$$

That is, *Noise(img)* is equal to the product of the three noise components.

The three noise components represent the "random" traffic variations at different time scales: *Noise*(*D*) for daily volumes, *Noise*(*H*) for hourly volumes, and *Noise*(*T*) for t^{dur}-minute volumes. Even though actual hourly volumes would be dependent on the daily volumes and the t^{dur} interval volumes would be dependent on the hourly volumes, the "random" traffic variation in the three time scales would appear to be mutually independent, since any of them does not provide any useful information about the "random" variation in the other two time scales. That is, knowing, for example, that the true daily volume V^D on a given day is greater than what is typically found on the day-of-week and month-of-year for that day, given true *AADT* (*i.e.*, $V^D(\delta) \times F^{MD}(m,d) > AADT$) would not provide information on how the true hourly volume $V^H(h,\delta)$ in a specified hour *h* relates the what would be expected in that hour, given the *true* daily volume. That is, it does not provide information on how $V^H(h,\delta)$ would compare to $V^D(\delta)/(24 \times F^H(h,d))$.

Since *Noise*(*img*) is a product of three independent noise components, the mean of *Noise*(*img*) would be the product of the means of the three noise components. E[*Noise*(*img*)] = E[*Noise*(*D*)] × E[*Noise*(*H*)] × E[*Noise*(*T*)]: Given the definitions of the three *Noise* components, we assume that the mean of each equals one. (Empirical analysis using ATR data supports that E[*Noise*(*D*)] and E[*Noise*(*H*)] are both approximately 1, and the analysis below argues for E[*Noise*(*T*)] = 1.) Therefore, E[*Noise*(*img*)] = E[*Noise*(*D*)]×E[*Noise*(*H*)]×E[*Noise*(*T*)] = 1×1×1 = 1. Since E[*Noise(img)*] = E[$AADT^{img} / AADT$] = E[$AADT^{img}$] / *AADT* = 1, it follows that E[$AADT^{img}$] = 1. Indeed, the empirical studies conducted to date (see and McCord *et al., (2003)* and Table1 in the text) support this result.

Since the three noise components are assumed to be independent and their means are equal to one, one can derive that the variance of the product of the three components is the sum of their variances plus all the variance cross-products That is, using σ_{img}, σ_D, σ_H, and σ_T, respectively, to denote the standard deviations of *Noise*(*img*), *Noise*(*D*), *Noise*(*H*), and *Noise*(*T*),

$$\sigma_{img}^2 = \sigma_D^2+\sigma_H^2+\sigma_T^2+\sigma_D^2\sigma_H^2+\sigma_H^2\sigma_T^2+\sigma_D^2\sigma_T^2+\sigma_D^2\sigma_H^2\sigma_T^2. \quad (A.5)$$

The values of the standard deviations can be shown to be small – on the order of 10^{-1}. Therefore, the variances will be on the order of 10^{-2}, and the cross-product terms on the order of 10^{-4} and 10^{-6}. We ignore these second order terms and approximate the variance of *Noise*(*img*) as

$$\sigma_{img}^2 \approx \sigma_D^2+\sigma_H^2+\sigma_T^2, \quad (A.6)$$

Equation (A.6) indicates that the variance of *Noise(img)* is approximately the sum of the variance of *Noise*(*D*), *Noise*(*H*) and *Noise*(*T*). Models such as the one we propose below for *Noise(T)* could conceivably be developed for *Noise(D)* and *Noise(H)*. However, the large quantity of daily and hourly

volumes that DOTs regularly collect and store across their statewide networks with ATRs allow direct estimation of these noise components. Specifically, the variances of the *Noise*(*D*) and *Noise*(*H*) terms can be estimated by the sample variances:

$$= \frac{}{} \quad ^2 \qquad \text{(A.7)}$$

and

$$= \frac{}{} \quad ^2 \qquad \text{(A.8)}$$

where and are the sample means for *Noise(D)* and *Noise(H)*, respectively, and *n* is the sample size.

Recall from Equation (A.3a) that *Noise(T)* equals the ratio of estimated hourly volume and the true hourly volume, where the estimated hourly volume is given by $\hat{V}^H(h,\delta) = N^{veh}/t^{dur}$. *Noise(T)* represents sub-hourly variation in traffic flows. DOTs would not regularly collect and store the data allowing for direct empirical estimates of this noise component. Moreover, sub-hourly temporal patterns are not expected to be as stable as the hourly or daily patterns. We are interested in estimating AADT from existing imagery collected in relatively free-flow conditions on uninterrupted freeways. To increase the likelihood of good lighting for the imagery, most air photos are taken around noon, and many freeway segments are near free-flow conditions. Moreover, if there are no recent accidents on the freeway, traffic would likely be in steady-state, non-transitioning conditions.

Under free-flow, steady-state conditions, traffic volumes would be reasonably modeled by a Poisson distribution. A realization N^{veh} of the random variable can be observed from an image, and probability *p* can be considered to be the proportion of the hour "observed" in the image, namely, t^{dur} [hrs]/1 [hr] = t^{dur}. In this way, $V^H(h, \delta)$ is the unknown parameter of interest. The mean and standard deviation of *Noise*(*T*) can then be derived from the negative binomial distribution.

$$(V^H(h,\delta) - N^{iveh} \mid \text{r} = N^{veh}, p = t^{dur}) \sim \textit{Negative}\ \text{Bin}\ (r = N^{veh}, p = t^{dur}) \qquad \text{(A.9)}$$

That is, given the observed number of vehicles N^{veh} from an image, the difference between unknown hourly volume $V^H(h, \delta)$ and observed volume N^{veh} follows a negative binomial (r, p) distribution, where using common negative binomial terminology, $r = N^{veh}$ is the number of success, $V^H(h, \delta)$ represents the number of "trials" to produce r successes, and $p = t^{dur}$ represents the "success probability".

We assume that the equivalent count interval t^{dur} determined from the images is the true equivalent count interval. We also ignore measurement errors in determining the number of vehicles in the imaged portion of the segment and the length of the imaged portion of the segment. Our emphasis is on the errors involved with converting measured data to an AADT estimate and not on the measurement errors associated with either imaging or ground-based technologies. With these assumptions, according

to (A.3a), and recalling the results for mean and standard deviation of a negative binomially distributed random variable ($V^H(h, \delta) - N^{veh}$ in this case), the mean and standard deviation of $V^H(h, \delta)$, can be derived, respectively, as

$$E(V^H(h, \delta)) = \frac{}{} = - = \frac{}{} \quad \text{(A.10a)}$$

$$\text{Var}(V^H(h, \delta)) = \frac{}{} \quad \text{(A.10b)}$$

and the mean and standard deviation of *Noise*(*T*), can then be derived, respectively, as

$$E[Noise(T)] = E\left[\frac{}{}\right] = -E\left[\frac{}{}\right] - \frac{}{} = 1 \quad \text{(A.11a)}$$

$$\text{Var [Noise(T)]} = = \text{var}\left[\frac{}{}\right] = \text{var}\left[\frac{}{}\right] = \text{var}\left[\frac{}{}\right] \quad \text{(A.11b)}$$

According to the delta method using second-order Taylor expansions to approximate the variance of a function of a random variable,

$$\text{var}\left[\frac{}{}\right] = \frac{}{} \quad \text{(A.12)}$$

and Equation (A.11b) will become

$$\text{*} \frac{}{} = \frac{}{} = \frac{}{} \quad \text{(A.13)}$$

Substituting from Equations (A.2b)

$$\frac{}{} \quad \text{(A.14)}$$