# Appendix D: Safety Analysis Methodology

## Previous Work on Applying Statistical Techniques in Regional Safety Modeling

Crash safety literature has produced a wide breadth of statistical techniques applied to variety of geographical entity when it comes to developing area-wide predictive crash models. Some of these previous studies are listed in a tabular format (Table 1) showing the types of spatial aggregation used, types of crashes modeled and modeling techniques adopted. This representation is not meant to be exhaustive and has excluded studies that modeled pedestrian and bicycle crashes deliberately to maintain relevancy to this study which models total and severe crash types only. The applied models, as shown in the Table, can be broadly characterized based on their spatial accountability in the model structure- thus, spatial and non-spatial models.

Among a wide spectrum of non-spatial models it can be observed from the literature that Negative Binomial (NB) model structure is favored particularly because of its ability to handle overdispersed crash data well. And the list of studies that applied NB model in crash predictions for different spatial units are quiet long (Amoros and Laumon, 2003; Noland and Oh, 2004; Hadayeghi et al., 2003, 2006, 2007; Aguero-Valverde and Jovanis, 2006; Quddus, 2008; Lord and Mannering, 2010; Naderan and Shahi, 2010; Abdel-Aty et al., 2011; Pirdavani et al., 2012; Karim et al., 2013; Pulugurtha et al., 2013). Since crashes are aggregated for a spatial entity, it is intuitive to consider the presence of spatial correlation in the model structure. But that increases the model complexity and data needs to some degree. Aguero-Valverde (2013) argued that spatial models, by dint of accounting spatial correlation, has potential to increase model fit by estimating 'pool strengths' from the spatial neighbors, and spatial effects can be surrogates for unknown and relevant covariates (Dubin, 1988; Cressie, 1993). As shown in Table 1, similar to NB, spatial models are being widely explored for predicting macro-level crashes and specifying a hierarchical Bayesian model that can account for overdispersion appears to be a popular technique among the researchers (Aguero-Valverde and Jovanis, 2006; Quddus, 2008; Huang et al., 2010; Karim et al., 2013; Aguero-Valverde, 2013). This study applied both NB and Bayesian hierarchical models for forecasting total and severe crashes. Two forms of Bayesian models were specified- one without accounting for spatial correlation (i.e., non-spatial Bayesian) and another accounting for spatial correlation. As such, there were three candidate models to compare 'classical versus Bayesian' and 'non-spatial versus spatial' modeling approaches.

The need for proactively forecasting safety for long range transportation plans has been reverberated in many of these studies. For example, Karim et al. (2013) evaluated the spatial effects of the occurrence of crashes in the Traffic Analysis Zones (TAZs) of Metro Vancouver to improve model fit and inference capability understanding that their effort will allow transportation authorities and planners to estimate safety proactively 'at a very early stage of transportation planning'. They (Karim et al., 2013) concluded that the spatial effects need to be considered in the crash prediction models to avoid any potential bias associated with model misspecification. Pulugurtha et al. (2013) estimated crashes for the TAZs of North Carolina based on land use characteristics and argued that their models can be used in safety conscious planning, land use decisions, and long range transportation plans. The authors (Pulugurtha et al., 2013) used NB models with a wide variety of land use covariates to model total, injury and property damage only (PDO) type crashes separately. Pirdavani et al. (2012) developed different zonal prediction models

for injury crashes and commented that the main purpose of their study was to develop planning-level predictive tool in order to evaluate safety for different travel demand management policies.
This report investigates forecasts found from multiple crash models developed for mid-region of New Mexico comprising Bernalillo, Sandoval, Valencia, and Torrance County. A methodical long range transportation planning technique, known as scenario planning, was utilized for developing different scenarios which were then used for forecasting safety. A twenty-five year planning horizon (2015-2040) was adopted for the study.

The rest of the report is structured as follows. The immediate next section describes the scenario planning process and the scenarios that were developed and applied in this study. The following section is about forecasting zonal parameters which illustrates the application of land use and travel demand models used for preparing datasets. The next two sections discuss crash models and parameter estimates. The section thereafter compares results from model forecasts for different planning scenarios. Finally the paper ends with a summary and concluding remarks.

# Scenario Planning

Since 2004, the Federal Highway Administration (FHWA) has encouraged transportation-focused scenario planning as an enhancement of the traditional transportation planning process. Scenario planning techniques are designed to help practitioners to consider how future changes in transportation, land use, demographics, or other factors could affect communities. At the core of scenario planning lies identifying land-use patterns as a dynamic variable affecting transportation networks, investments, and operations. Other potential variables may include demographic, economic, political, and environmental trends. These variables are used to develop alternated 'possibilities' or 'scenarios' that help stakeholders to understand how a region might look and function in the future, and make decisions for the present and prepare for future needs. (FHWA Scenario Planning Guidebook, February, 2011)

Three alternative scenarios were developed in this study. Each scenario is briefly described below.

## Alternative 1 (Trend Scenario)

This scenario continues the patterns from the early 2000s in which residential development was focused on single family housing in more peripheral parts of the region. This scenario assumes that commercial development is scattered across the region rather than in targeted centers. About half of new jobs, but about three-quarters of new housing are located west of the river (Rio Grande). Private vehicle travel remains the dominant mode for the vast majority of residents in the region. This scenario resembles continuing historical trend. Major scenario components include-

- Low and medium-density residential housing in previously undeveloped areas
- No particular emphasis on mixed-use development or along transit corridors
- Commercial development is scattered around region rather than concentrated in particular areas
- Assumes continued reliance on private vehicles for most trips.

### Alternative 2 (Preferred Scenario)

This scenario reflects a range of trends in housing preferences and travel behavior across the region. Parcels within a ½-mile radius of existing and future transit stops were designated for medium-density mixed-use development and multi-family, and those within a ¼-mile radius were designated for high-density mixed-use. Emphasis is placed on compact development in targeted locations near transit to meet the demands of a range of age demographics. An increased preference for alternative modes and increased spending on public transportation was emphasized. Major scenario components include-

- Development on activity centers and corridors near premium transit
- Accessory dwelling units to meet senior and multi-generational housing needs
- Multi-family housing near transit
- Greater emphasis on mixed-use development
- More transportation options and increased preference for proximity to services and entertainment.

### Alternative 3 (Preferred Constrained Scenario)

This scenario reflects all the major components of Alternative 2 except a constraint was imposed on the road network. Each of these scenarios were developed for 2015-2040 forecast years. Unlike Alternate 2, this scenario restricted the growth of highway and transit network at year 2025. As such all growth beyond 2025 would be based on constrained network capacities. More detailed discussion on the networks are provided in the Travel Demand Model section.

## Forecasting Zonal Parameters

The study utilizes 914 data analysis subzones (DASZ) which are geographic entities similar to traffic analysis zones (TAZ). These zones contain the entirety of four counties of New Mexico (Bernalillo, Sandoval, Valencia and Torrance County) comprising the Albuquerque metropolitan area.
Figure 1 shows the study area and its relative location with respect to the state of New Mexico. A few DASZs (north of Torrance County and east of Bernalillo County) that were outside the County boundary (in
Figure 1) were part of Santa Fe County. Crashes that took place between 2006 and 2010 were analyzed. Aggregated total crashes and severe crashes per DASZ were modeled using 2010 socioeconomic (SE) data as the independent variables. Severe crashes were defined as the sum of fatal and injury type crashes. Bernalillo is an urban County and captures the largest share of crashes in the study region. Valencia, Torrance and parts of Santa Fe County in the study area are mostly rural. About 88% of total crashes and about 86% of severe crashes took place in Bernalillo County alone. Together, Bernalillo and Sandoval County captured 95.3% of total crashes and 94.7% of severe crashes in the study area. Severe crashes were about 30% of the total crashes region-wide.

For each scenario a land use model was specified. The land use model iteratively ran with two travel demand models. Land use models took a wide array of observed variables considering 2012 as the base year. This model then ran till 2040 forecast year. The following sub-sections discuss more on each of these models.

## Land Use Model and Travel Demand Model

The Open Platform for Urban Simulation (OPUS) was used for land use modeling. OPUS architecture is primarily based on the UrbanSim project (The Open Platform for Urban Simulation and UrbanSim Version 4.3, January, 2011). UrbanSim is a software-based simulation system that incorporates the interactions between land use, transportation, the economy, and the environment. It supports planning and analysis of urban development and helps explore the effects of infrastructure and policy choices on community outcomes such as motorized and non-motorized accessibility, housing affordability, etc. More on UrbanSim can be found at urbanism.org. Currently multiple planning organizations in USA have adopted UrbanSim for operational planning use; examples include- Maricopa Association of Governments, Metropolitan Transportation Commission, Puget Sound Regional Council, etc.

The land use model utilized parcels as the smallest geographic entity for analyzing, aggregating, and creating the database structure. The OPUS base year database was developed using 2012 data which was used as initial inputs for starting each scenario simulation. Zonings for each of the alternative scenarios were input to OPUS. OPUS then produced a socioeconomic forecast for every five years between 2015 and 2040.

The travel demand model (TDM) was built in Citilabs Cube 6.1.0. Table 2 and Table 3 show different networks that were used for different planning scenarios. Unlike the land use model, the travel model scenarios were built for years 2025 and 2040 only. Therefore, the travel time skim was fed from travel demand model to OPUS in 2025 and 2040 only. As mentioned earlier, Alternative 3 represents a constrained network scenario. For 2025, the Alternative 3 roadway and transit networks were constrained to their corresponding 2012 networks (without any improvement). For 2040, the Alternative 3 roadway and transit networks were restricted to their corresponding 2025 networks.

## Iterative Modeling between OPUS and TDM

Simulation of each scenario was started in UrbanSim. UrbanSim was interfaced with Cube at years 2025 and 2040 where UrbanSim 'called' Cube to perform a TDM analysis to predict travel conditions for those years. Therefore, each of the scenario-runs constituted two TDMs for years 2025 and 2040. Figure 2 depicts the exchange of data that took place in the iterative process between OPUS and TDM. Land use predictions from UrbanSim got input to the TDM, and travel conditions were input to the subsequent annual iterations of the UrbanSim land use model system. When UrbanSim is connected to TDM, it generates a summary of the household and job data at the DASZ level which feeds into TDM as an input data.

Because of the loop-back between OPUS and TDM, the forecasted socioeconomic variables are thought to account for the future effect of transportation infrastructure. Therefore, it would be redundant to calibrate safety models with both socioeconomic and transportation-related variables. Moreover, as land use and transportation are proven to influence one-another, the dynamics of feeding socioeconomic data from OPUS into TDM (for a simulation year) would have impact on forecasted travel time (output from Cube); and in a similar way feeding travel skims from TDM to OPUS would have its influence on the socioeconomic data for the future years.

The following section provides descriptive statistics of the socioeconomic variables that were used in both calibration and forecasting of the regional safety models.

## Definition and Descriptive Statistics of the Zonal Variables

Table 4 defines each of the variables and provides their descriptive statistics. Each of the variables were aggregated at the DASZ level. Some of the variables were transformed to minimize heteroscedasticity of their variance. Total number of employed people per zone were divided into three categories- basic, retail and service. Basic employment being jobs related to agriculture and manufacturing industry, retail employment captured number of people working in the retail sector, and the number of people in service were defined in the service employment variable. It is important to note that these employment variables signifies the number of people who are employed in a DASZ, and they may or may not be a resident to that DASZ.

# Regional Crash Prediction Models

Crashes were modeled using a Negative Binomial model and two Bayesian models. Bayesian models were given a Poisson-Lognormal structure to fit the crash data appropriately. The difference between two Bayesian models lied in incorporating spatial heterogeneity. One of the Bayesian models, which will be termed as 'spatial Bayesian model', had a spatial error component defined in the model structure. The other Bayesian model did not have any spatial error component and will be termed as 'non-spatial Bayesian model'. Previous studies (El-Basyouny and Sayed, 2009; Huang and Abdel-Aty, 2010; Siddiqui and Abdel-Aty, 2012) have shown that Bayesian models with spatial error component tend to fit and predict crash data well. However, the study investigated these models not only from their strict predictive fits but also in regards to their crash forecasts for future planning years.

Negative Binomial models are relatively easy to estimate especially with built-in procedure available in a handful of commercial and open source statistical software. Bayesian models, on the other hand, can be a little bit of work in terms of coding and specifying an appropriate data structure. These models are possible to fit using open source software like R (The R Project) or OpenBUGS (openbugs.net); however, they demand that the modeler have a relatively greater degree of knowledge in coding. Also, to incorporate spatial weight matrix into the model, the modelers have to use some kind of mapping software (such as ArcMap). This study attempts to investigate if all this extra work indeed makes a difference when it comes to forecasting long range safety.

Each of these models is discussed below.

## Negative Binomial Model

It is the most simple among the three techniques applied in this study. NB regression is a type of generalized linear model in which the response variable is a count of the number of times an event occurs which in this case is 'occurrence of crashes'. The probability distribution of the response variable *y* can be given by (Hilbe, 2011):

$$P(y) = P(Y = y) = \frac{\Gamma(y + 1/\alpha)}{\Gamma(y + 1)\Gamma(1/\alpha)} \left(\frac{1}{1 + \alpha\mu}\right)^{1/\alpha} \left(\frac{\alpha\mu}{1 + \alpha\mu}\right)^{y}$$

where, $\mu > 0$ is the mean of Y, and $\alpha > 0$ is the heterogeneity parameter.

## Non-spatial Bayesian Poisson-Lognormal Model

A Poisson-lognormal model was specified as follows:
$y[i] \sim \text{Poisson}\ (\mu[i])$
$\log(\mu[i]) = \beta_0 + \beta X_i + \theta[i]$
$\theta[i] \sim \text{Normal}\ (0, \tau_\theta)$
where,
$\beta_0$ = intercept term,
$\beta$'s are the coefficient estimates of the model covariates ($X_i$),
$\theta[i]$ = error component of the model capturing unstructured over-dispersion or unobserved heterogeneity component of the model, and
$\tau_\theta$ = precision parameter which is inverse of the variance; a prior gamma distribution is specified to $\tau_\theta$.

The variance ($1/\tau_\theta$) provides the amount of variation not explained by the Poisson assumption (Lawson et al., 2003). A uniform prior distribution was assumed for $\beta_0$. The model was run considering a non-informative Normal(0, 100000) prior for β's.

## Spatial Bayesian Poisson-Lognormal Model

The spatial Bayesian Poisson-Lognormal model included an explicit error component, ($\phi[i]$), to account for the portion of heterogeneity occurring due to spatial correlation. Spatial distribution was implemented by specifying an intrinsic Gaussian Conditional Autoregressive (CAR) prior with Normal($\overline{\phi}[i]$, $\tau_i^2$) distribution recommended by Besag (1974).

($\phi[i]$) is defined as-

$$\overline{\phi}[i] = \frac{\sum_{i\neq j} \phi[j] * W_{ij}}{\sum_{i\neq j} W_{ij}}$$

where, $W_{ij}$ is the element of adjacency matrix with a value of 1 if *i* and *j* are adjacent or 0 otherwise.

## Comparison of Model Fit

Both Bayesian models were initialized using non-informative priors for the intercept, β's, and error components. Each model had three Markov chains. Burn-in sample size and 'thinning' was set to 5000 and 5, respectively. Model convergence and performance were tested based on chain convergence (trace plots), density plots, and Brooks-Gelman-Rubin statistics. OpenBUGS provides Bayesian Credible Intervals (BCIs) to draw inference on the significance of the parameter estimates.

For classical models such as Negative Binomial, Akaike Information Criterion (AIC) is used for comparing non-nested models. AIC is defined as-

$$\text{AIC} = -2\log\left(p\left(y|\hat{\theta}\right)\right) + 2k = D\left(\hat{\theta}\right) + 2p$$

where, $\hat{\theta}$ = maximum likelihood estimate, and p = number of parameters in the model.
The term 2p in the above equation serve to penalize more complex model. For model comparison predictive ability forms a natural criterion. AIC fits well in this respect since it is designed to optimize predictions on a replicate dataset of the same size. And a model with a lower AIC is favored. (Lunn et al., 2012)

In a Bayesian context the posterior mean deviance $\overline{\mathrm{D}} = E[\mathrm{D}]$ has been suggested as a measure of fit. However, in analogy to AIC a measure of 'model complexity' is necessary to trade off against $\overline{\mathrm{D}}$ as more complex Bayesian model will fit the data better- hence decreasing the value of $\overline{\mathrm{D}}$. To suggest a measure of effective number of parameters ($\mathrm{p_D}$ ) Spiegelhalter et al. (2002) used an informal information-theoretic argument defined by

$$\mathrm{p_D} = \mathrm{E}_{\theta|\mathrm{y}}\left[-2\log\left(\mathrm{p}(\mathrm{y}|\theta)\right)\right] + 2\log\left(\mathrm{p}\left(\mathrm{y}\middle|\tilde{\theta}(\mathrm{y})\right)\right) = \overline{\mathrm{D}} - \mathrm{D}(\tilde{\theta})$$

where, $\tilde{\theta}$ is a 'good' plug in estimate of θ. If we consider $\tilde{\theta} = \mathrm{E}[\theta|\mathrm{y}] = \bar{\theta}$, $\mathrm{p_D}$ reduces to 'posterior mean deviance' minus 'deviance of posterior means'. For large sample size or in presence of non-informative or 'vague' prior (which is the case in this study), when conditions for asymptotic normality is present, $\bar{\theta} \approx \hat{\theta}$, the maximum likelihood estimate and $\mathrm{p_D}$ reduces to p, the total number of parameters in the model. For complete discussion on this issue readers are referred to Lunn et al., 2012.

The measure of $\overline{\mathrm{D}}$ can now be combined with model complexity parameter $\mathrm{p_D}$ to calculate an AIC-like measure called Deviance Information Criterion (DIC):

$$\mathrm{DIC} = \overline{\mathrm{D}} + \mathrm{p_D} = \mathrm{D}(\bar{\theta}) + 2\mathrm{p_D}$$

DIC thus acts as a generalization of AIC. Since for non-informative prior $\bar{\theta} \approx \hat{\theta}$, this results $\mathrm{p_D} \approx \mathrm{p}$ and DIC ≈ AIC.

It was found that for total crash estimation (Table 5) the Negative Binomial, non-spatial, and spatial Bayesian model had the following values respectively: 9355 (AIC), 6637 (DIC), and 5788 (DIC). These values for severe crash estimation (Table 6) were 7273 (AIC), 5533 (DIC), and 1083 (DIC) for Negative Binomial, non-spatial, and spatial Bayesian models, respectively. The difference among DIC/AIC values are considerably large which signifies the superiority of spatial Bayesian model in terms of predictive fit for both total and severe crashes.

## Parameter Estimates

Table 5 and Table 6 provides parameter estimates for total and severe crash models, respectively. For both total and severe crashes total number of signalized intersections, population count, and employment types (basic, retail, and service) were found to be positively associated. These associations are intuitive have been supported by previous studies (Quddus, 2008; Pirdavani et al., 2012; Aguero-Valverde, 2013).

Median income was negatively associated with both types of crashes indicating that poverty stricken zones are more prone to crashes. This association is also concurrent with previous findings (Noland and Quddus, 2004b; Aguero-Valverde and Jovanis, 2006; Huang et al., 2010; Pirdavani et al., 2012). For Albuquerque metropolitan area this finding may be particularly important to ponder upon as poverty rate in Albuquerque has steadily increased over the past seven years (City of Albuquerque Progress Report).

The number of single family dwelling unit (SFDU) was consistently negatively associated with total and severe crashes in both NB and non-spatial Bayesian models. On contrary, the number of multiple family dwelling unit (MFDU) was positively associated with both crash types in all three models. However, spatial Bayesian model for severe crashes deemed estimate of MFDU as statistically not significantly different from zero at 95% Bayesian Credible Interval (BCI). Also, for both crash models spatial Bayesian model had positive estimates for SFDU.

Among the County dummy variables, only Bernalillo and Valencia County dummies were found to be significantly different from zero at 95% BCI. Both of these dummy variables were negatively associated with total and severe crash types.

## Evaluation of Regional Safety Models for Different Planning Scenarios

All three of the above mentioned models were applied to generate forecasts of total and severe crashes in every five-year interval starting 2011 until 2040. The forecasted socioeconomic data from land use and travel demand model iterations were used to calculate crash estimates for the future years. Table 7 provides the percent increase of crashes for the overall study area with respect to the base case (crashes occurring in between 2006 and 2010 inclusive) for three alternative scenarios.

For all three alternative scenarios crash forecasts from the Bayesian spatial model were minimum among three candidate models. Considering percent increase of crashes over the planning horizon, Alternate 1 scenario was found to be the safest for both total and severe crash types. Crash forecasts between Alternative 2 and 3 showed slight differences. Recall that the difference between Alternate 2 and 3 were in terms of constrained roadway and transit networks for years 2025 and 2040 (Table 2 and Table 3) which is why these two scenarios have the same forecasts until 2025. It is possible that the difference in socioeconomic forecasts between Alternate 2 and 3 were not large enough to be reflected in their corresponding crash forecasts from 2025 till 2040.

As found in the previous section, the spatial Bayesian model performed the best in terms of crash predictability. Its superior goodness of fit most likely lies in being able to capture spatial heterogeneity among the DASZs. It was found that for total crashes, about 77.5% of the error was captured by the spatial error term ($\phi[i]$)). The same for severe crashes was about 77.1%. Inclusion of explicit error component for spatial heterogeneity seems more practical while modeling spatially aggregated count data. However it increases model complexity to some degree. If predictive fit of these models is put aside and only relative safety forecasts are compared among three scenarios, it can be observed that all three candidate models points towards Alternate 1 scenario to be safest in the planning horizon irrespective of modeling techniques. This implies that the approach of a safety forecast in the short-term versus long-term can be dictated by the accuracy of the predictability needed/expected from a regional safety model. Also, fitting spatial Bayesian models up until now in OpenBUGS or R involves multiple steps of data preparation and a certain level of coding expertise; and therefore, is not as straight forward as fitting NB models.

Similar to Aguero-Valverde and Jovanis, 2006 and Siddiqui et al., 2012 this study found that the non-spatial Bayesian models have better goodness of fits than that of NB models for predicting crashes. The forecasts for both total and severe crashes from the non-spatial Bayesian model, however, provided the largest variations between 2011 and 2040.

To further investigate the forecast pattern of the alternative scenarios the resolution was changed from the total modeling area to the County level. Table 8 lists the forecasted total and severe crashes for three scenarios based on the spatial Bayesian model. The forecasts for Bernalillo County were much similar to the overall safety forecast. And this is not surprising since Bernalillo County takes the lion share of the crashes in the modeling area. However, Sandoval County showed some interesting results. Sandoval County takes the second largest share of total and severe crashes in the study area (7.4% and 8.5% respectively). It was found that Alternate 2 was the safest in Sandoval County for both total and severe crashes. This indicates that the spatial aggregation can play an important role in the decision making process. In general Sandoval County was more 'responsive' to the crashes. In thirty years total number of crashes in Sandoval County increased by almost three times in Alternative 1, and about two times for Alternative 2 and 3 while compared to that of Bernalillo County. Although, severe crashes in Sandoval County did not increase by the scale of total crash increase, the County still experienced considerably higher percentages of severe crashes compared to its neighboring Bernalillo County.

As mentioned before, Bernalillo and Sandoval counties captured the majority of crashes. And from the above analysis certain differences in crash forecasts were found in these two counties. Therefore, to have a better understanding about County specific crash forecast, and also to gain insight into model transferability (County-specific versus region-wide), total and severe type crashes were modeled separately specific to Bernalillo and Sandoval County- as such four spatial Bayesian models were developed and are presented in Table 9 and Table 10. Out of 914 DASZs in the study area, Bernalillo County comprised of 660 DASZs, and there are 142 DASZs in Sandoval County. The forecasts from these models are presented in Table 12. Some of the interesting observations from the County-specific analysis are discussed below.

Apart from the range of parameter estimates the main difference between Bernalillo and Sandoval total crash models (model-8a and -9a) were in the signs of coefficient estimates of total population, median income and single family dwelling units (SFDU). The negative estimate for population and positive estimate for median income in Bernalillo County (model-8a) is counter intuitive. However, positive association between number of crashes and affluent areas are not quiet uncommon in spatial crash modeling. Aguero-Valverde (2013) found similar association and reported that percentage of person under poverty line living in cantons (smaller political units with a local government) of Costa Rica had lower crash frequency for injury and property damage only types of crashes. The direction of association of parameters in spatial crash modeling can be affected by the size (scale) of the modeling-area, spatial unit of aggregation, and the confounding effect among the parameters. The correlation matrices for the County-specific models were provided in Table 11. For all four models (model-8a, 8b, 9a, 9b) moderate to high negative correlations were observed between total population & median income, and total population & SFDU.

While comparing between region-wide (Table 8) and County-specific (Table 12) total crash forecasts, both Bernalillo and Sandoval County followed similar pattern. Similar to the region-wide model, Alternative 1 scenario was found to be the safest in Bernalillo County (with respect to total crashes). Alternative 2 was found to be the safest for total crashes in Sandoval County. Once again, County forecasts from Alternative 2 and 3 were found to be very similar. Interestingly, the forecasts during the end of thirty year planning horizon were also found to be close and considerably similar- 2036-2040 forecast between Alternative 1 and 2 differed about 5% and 3% in Bernalillo and Sandoval County, respectively. But the change in the percent increase of total crash between 2011-2015 and 2036-2040 were higher in Sandoval County than in Bernalillo County.

Similar to the total crash models (model-8a and -9a) the differences in the signs of parameter estimates for total population and median income were observed for Bernalillo and Sandoval County's severe crash models (model-8b and -9b) as well. But SFDU was positively associated with the number of severe crashes in both of the counties. In addition, a similar pattern in the increase of severe crashes was observed between the region-wide (Table 8) and County-specific (Table 12) models. In regards to severe crashes Alternative 1 scenario was found to be the safest for Bernalillo County, and Alternative 2 was preferred for Sandoval County.

All covariates presented in Table 9 and Table 10 were statistically significant at 95% Bayesian Credible Interval. The contribution of spatial heterogeneity within the total error structure was less in Sandoval County than in Bernalillo County for both total and severe crash models. ϕ[i] for Sandoval County was 24.6% and 16.3% for total and severe crashes, respectively. The same (ϕ[i]) for Bernalillo County was 80.8% and 80.1% for total and severe crashes, respectively.

Given that the County-specific analysis did not necessarily provide any different pattern in the forecast for both total and severe crashes when compared with the region-wide model, it is reasonable to conclude that a region-wide model would be a more practical and less computationally-intensive route for long range safety forecasts. However, particular differences in parameter estimates were observed between the county-specific models. Therefore it might be beneficial and worth estimating County-specific models for a relatively short-term safety predictions. Also, as commented before, the differences between different scenario forecasts were found to be wider in near future than distant years.

## Summary

This study attempted to evaluate forecasting performances of regional safety models for different alternative scenario planning. Regional crash prediction models are often compared strictly based on their predictive fit. This study compared models beyond their predictive performance and investigated the role of model complexity (modeling techniques) and model granularity (spatial aggregation) may have in improving long range planning forecasts.

The study used 2012 as a base year for the independent parameter set. Demographic, socioeconomic, roadway and transit networks were utilized in preparing the base year dataset. These parameters were used to model total and severe crashes that occurred in between 2006 and 2010 inclusive, in the study area which comprised the entirety of four counties in Central New Mexico. In order to forecast exogenous variables for future years, UrbanSim (a land use model) and Cube (travel demand model) were iteratively run for each of the alternative scenarios. The socioeconomic variables were forecasted every five years between 2011 and 2040. Forecasted socioeconomic data were then used to forecast crashes for every five years in the same time span.

Among the three candidate models fitted for both total and severe crashes, the Bayesian model accounting for the spatial heterogeneity among DASZs outperformed the Negative Binomial and Bayesian model that did not account for spatial error in terms predictive fit. In general, the spatial Bayesian model forecasted the smallest increase in crash occurrences in future years. This may be a particularly important finding in terms of applying a model with a better fit since an unreasonably high increase of crashes (thus, deteriorating safety) would adversely affect public perception in the scenario planning process. The widest band of increase in crashes was observed from non-spatial Bayesian models which, in spite of its better predictive fit than Negative Binomial models, provided similar or worse safety forecasts for the alternative scenarios.

The effect of spatial granularity (region-wide versus County-specific) on model estimates and in turn on safety forecast was investigated. In order to understand whether an overall forecast would reveal a similar forecast pattern when compared to a smaller scale, County-specific models were developed for total and severe crashes. Percent increase in total and severe crashes showed a similar pattern (both in direction and scale) when compared between the region-wide and county-specific models.

## REFERENCES


1. Abdel-Aty, M., C. Siddiqui, H. Huang, and X. Wang. Integrating Trip and Roadway Characteristics to Manage Safety in Traffic Analysis Zones. In Transportation Research Record: Journal of the Transportation Research Board, No. 2213, 2011, pp. 20–28
2. Aguero-Valverde, J., Jovanis, P. P., 2006. Spatial Analysis of Fatal and Injury Crashes in Pennsylvania. Accident Analysis and Prevention 38, pp. 618-625
3. Aguero-Valverde, J., 2013. Multivariate Spatial Models of Excess Crash Frequency at Area Level: Case of Costa Rica. Accident Analysis and Prevention 59, pp. 365-373
4. Amoros, E., J. L. Martin, and B. Laumon. Comparison of Road Crashes Incidence and Severity Between Some French Counties. Accident Analysis and Prevention, Vol. 35, No. 4, 2003, pp. 537–547
5. Besag, J., 1974. Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical Society B 36(2), pp. 192-236
6. City of Albuquerque Progress Report, Residents Living in Poverty, http://www.cabq.gov/progress/human-and-family-development/dc-10/indicator-10-1, Accessed on October 24, 2014
7. Cressie, N. A. C., 1993. Statistics for Spatial Data. John Wiley & Sons, New York
8. De Guevara, F. L., Washington, S. P., and Oh, J., 2004. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. Transportation Research Record, No. 1897, pp. 191-199.
9. Dubin, R. A., 1988. Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms. Review of Economics and Statistics 70 (3), pp. 466-474
10. El-Basyouny, K., and Sayed, T., 2009. Collision Prediction Models using Multivariate Poisson-Lognormal Regression. Accident Analysis and Prevention 41, pp. 820-828
11. FHWA Scenario Planning Guidebook, U.S. Department of Transportation, February 2011
12. Hadayeghi, A., Shalaby, A. S., and Persaud, B. N., 2003. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. Transportation Research Record, No. 1840, pp. 87-95.
13. Hadayeghi, A., A. S. Shalaby, B. N. Persaud, and C. Cheung. Temporal Transferability and Updating of Zonal Level Accident Prediction Models. Accident Analysis and Prevention, Vol. 38, No. 3, 2006, pp. 579–589
14. Hadayeghi, A., A. S. Shalaby, and B. N. Persaud. Safety Prediction Models: Proactive Tool for Safety Evaluation in Urban Transportation Planning Applications. In Transportation Research Record: Journal of the Transportation Research Board, No. 2019, 2007, pp. 225–236
15. Hilbe, J., 2011. Negative Binomial Regression, 2nd Edition, New York: Cambridge University Press
16. http://www.openbugs.net/w/FrontPage , Accessed on October 20, 2014
17. Huang, H., and Abdel-Aty, M., 2010. Multilevel Data and Bayesian Analysis in Traffic Safety. Accident Analysis and Prevention 42, pp. 1556-1565

18. Huang, H., Abdel-Aty, M. A., and Darwiche, A. L., 2010. County-level Crash Risk Analysis in Florida: Bayesian Spatial Modeling. Transportation Research Record 2148, pp. 27-37
19. Karim, M. A., Wahba, M. M., and Sayed, T. Spatial Effects on Zone-Level Collison Prediction Models. Transportation Research Record: Journal of the Transportation Research Board, No. 2398, pp. 50-59
20. Lawson, A. B., Browne, W. J., and Rodiero, C. L. V, 2003. Disease Mapping with WinBUGS and MLwiN. John Wiley & Sons Ltd., England
21. Lord, D., and F. Mannering. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. Transportation Research Part A: Policy and Practice, Vol. 44, No. 5, 2010, pp. 291–305.
22. Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D., The BUGS Book – A Practical Introduction to Bayesian Analysis, Chapman and Hall/CRC, October 2012
23. Naderan, A., and J. Shahi. Aggregate Crash Prediction Models: Introducing Crash Generation Concept. Accident Analysis and Prevention, Vol. 42, No. 1, 2010, pp. 339–346
24. Noland, R. B., and L. Oh. The Effect of Infrastructure and Demographic Change on Traffic-Related Fatalities and Crashes: A Case Study of Illinois County-Level Data. Accident Analysis and Prevention, Vol. 36, No. 4, 2004, pp. 525–532
25. Noland, R. B., and Quddus, M. A., 2004a. Analysis of Pedestrian and Bicycle Casualties with Regional Panel Data. Transportation Research Record 1897, pp. 28-33.
26. Noland, R. B., and Quddus, M. A. A., 2004b. A Spatially Disaggregate Analysis of Road Casualties in England. Accident Analysis and Prevention 36, pp. 973-984.
27. Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., and Wets, G., 2012. Application of Different Exposure Measures in Development of Planning-Level Zonal Crash Prediction Models. Transportation Research Record: Journal of the Transportation Research Board, No. 2280, pp. 145-153
28. Pulugurtha, S. S., Duddu, V. R., and Kotagiri, Y., 2013. Traffic Analysis Zone Level Crash Estimation Models Based on Land Use Characteristics. Accident Analysis and Prevention 50, pp. 678-687
29. Quddus, M. A., 2008. Modelling Area-wide Count Outcomes with Spatial Correlation and Heterogeneity: An Analysis of London Crash Data. Accident Analysis and Prevention 40, pp. 1486-1497
30. Siddiqui, C., and Abdel-Aty, M., 2012. Nature of Modeling Boundary Pedestrian Crashes at Zones. Transportation Research Record: Journal of the Transportation Research Board, No. 2299, pp. 31-40
31. Siddiqui, C., Abdel-Aty, M., and Choi, K., 2012. Macroscopic Spatial Analysis of Pedestrian and Bicycle Crashes. Accident Analysis and Prevention 45, pp. 382-391
32. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Linde, A. V. D., 2002. Bayesian Measures of Model Complexity and Fit, Journal of the Royal Statistical Society B 64 (4), pp. 583-639
33. The Open Platform for Urban Simulation and UrbanSim Version 4.3, User Guide and Reference Manual, The UrbanSim Project, University of California Berkley, and University of Washington, January 23, 2011
34. The R Project for Statistical Computing, http://www.r-project.org/, Accessed on October 20, 2014
35. www.urbanism.org, Accessed on October 20, 2014

# TABLES

**Table 1: Few Examples of Previous Studies on Macro-level Crash Prediction**

| Reference | Spatial Aggregation | Applied Models | Modeled Crash Type(s) |
|---|---|---|---|
| Hadayeghi et al. (2003) | Traffic Zones of City of Toronto | Negative Binomial; Geographically Weighted Regression | Total; Severe (fatal and nonfatal injury) |
| Aguero-Valverde and Jovanis (2006) | Counties of Pennsylvania | Negative Binomial; Full Bayes Hierarchical Model with Spatial and Temporal Effects | Injury; Fatal |
| Quddus (2008) | Census Wards of the Greater London Metropolitan Area | Negative Binomial; Spatial Autoregressive; Spatial Error Model; Spatial Poisson-Lognormal | Fatal; Serious Injury; Slight Injury |
| Huang et al. (2010) | Counties of Florida | Spatial Poisson-Lognormal | Total; Severe |
| Pirdavani et al. (2012) | Traffic Analysis Zones of Flanders, Belgium | Negative Binomial | Injury |
| Aguero-Valverde (2013) | Cantons of Costa Rica | Multivariate Spatial Model using Full Bayes Hierarchical Approach | Fatal; Injury; Property Damage Only |
| Karim et al. (2013) | Traffic Analysis Zones of Metro Vancouver, Canada | Negative Binomial; Spatial Poisson-Gamma | Total; Severe; Property Damage Only |
| Pulugurtha et al. (2013) | Traffic Analysis Zones from the City of Charlotte and Mecklenburg County, North Carolina | Negative Binomial | Total; Injury; Property Damage Only |

**Table 2: TDM Scenario Specifications for 2025**

| Scenario Parameters | Alternate 1 | Alternate 2 | Alternate 3 |
|---|---|---|---|
| Roadway Network | 2025 Network | 2025 Network | 2012 Network |
| Transit Network | 2012 Network (no improvement) | 2025 Network | 2012 Network (no improvement) |
| Socioeconomic (SE) Data | 2025 SE | 2025 SE | 2025 SE |

**Table 3: TDM Scenario Specifications for 2040**

| Scenario Parameters | Alternate 1 | Alternate 2 | Alternate 3 |
|---|---|---|---|
| Roadway Network | 2040 Network | 2040 Network | 2025 Network |
| Transit Network | 2012 Network + limited improvement | 2040 Network | 2025 Network |
| Socioeconomic (SE) Data | 2040 SE | 2040 SE | 2040 SE |

**Table 4: Variables Definitions**

| Variable Acronym | Variable Definition | Mean | Standard Deviation | Min | Max |
|---|---|---|---|---|---|
| *Response Variables* | | | | | |
| Cr06to10 | Total number of crashes between 2006 and 2010 inclusive | 94.06 | 124.38 | 0 | 1013 |
| SevCr06to10 | Total number of severe crashes between 2006 and 2010 inclusive. Severe crashes are sum of fatal and injury type crashes. | 27.66 | 36.07 | 0 | 275 |
| *Dependent Variables* | | | | | |
| SigInt | Total number of signalized intersections | 1.18 | 1.47 | 0 | 9 |
| SFDU | Single family dwelling units | 332.6 | 342.13 | 0 | 2180 |
| MFDU | Multiple family dwelling units | 81.8 | 202.73 | 0 | 2085 |
| CountyXXX | Dummy variables. Each for a County in the study area. Examples include- CountyBER for Bernalillo County, CountyVAL for Valencia County, etc. | - | - | - | - |
| LnPop | Logarithmic transformation of the total population ln(total population + 1) | 5.63 | 2.5251 | 0 | 8.75 |
| LnBasic | Logarithmic transformation of the Basic Employment ln(basic employment + 1) | 3.04 | 1.8557 | 0 | 10.04 |
| LnRetail | Logarithmic transformation of the Retail Employment ln(retail employment + 1) | 2.75 | 2.1598 | 0 | 7.6 |
| LnService | Logarithmic transformation of the Service Employment ln(service employment + 1) | 3.99 | 2.108 | 0 | 9.35 |
| LnEmp | Logarithmic transformation of the Total Employment which is a sum of Basic, Retail, and Service Employment ln(total employment + 1) | 4.71 | 2.1512 | 0 | 10.04 |
| LnMedInc | Logarithmic transformation of Median Income | 8.88 | 4.174 | 0 | 12.04 |
| Rent | Total number of rented units | 125.03 | 0.02 | 0 | 1819 |
| Own | Total number of owner occupied housing units | 259.35 | 0.0275 | 0 | 1751 |
| UNMenroll | Total number of students enrolled in University of New Mexico | 30.78 | 771.67 | 0 | 23111 |
| CNMenroll | Total number of students enrolled in Central New Mexico | 28.86 | 438.18 | 0 | 10944 |

**Table 5: Parameter Estimates from Total Crash Models**

| Variables | Non-Bayesian NB Model (Model-5a) | | Bayesian model without accounting spatial correlation (Model-5b) | | | | Bayesian model accounting spatial correlation (Model-5c) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Estimates | P-value | Mean | Std. Dev. | Bayesian Credible Interval | | Mean | Std. Dev. | Bayesian Credible Interval | |
| | | | | | 2.5% | 97.5% | | | 2.5% | 97.5% |
| SigInt | 0.2044 | < 0.001 | 0.2344 | 0.0302 | 0.1804 | 0.297 | 0.2028 | 0.0274 | 0.1529 | 0.261 |
| LnPop | 0.1131 | 0.0002 | 0.1203 | 0.032 | 0.0536 | 0.1709 | 0.0668 | 0.0318 | 0.0111 | 0.1265 |
| LnBasic | 0.1033 | < 0.001 | 0.1087 | 0.0261 | 0.0603 | 0.1583 | 0.0606 | 0.0265 | 0.0113 | 0.115 |
| LnRetail | 0.1511 | < 0.001 | 0.1994 | 0.029 | 0.1412 | 0.2529 | 0.124 | 0.0257 | 0.0716 | 0.1706 |
| LnService | 0.1619 | < 0.001 | 0.2237 | 0.0254 | 0.1748 | 0.2715 | 0.1781 | 0.0281 | 0.1254 | 0.2337 |
| LnMedInc | - 0.0641 | < 0.001 | - 0.0357 | 0.0125 | - 0.0595 | - 0.01042 | - 0.021 | 0.0197 | -0.0582 | 0.01493 |
| SFDU | - 1.951E-4 | 0.1748 | - 1.09E-4 | 1.66E-4 | - 4.29E-4 | 2.03E-04 | 1.59E-4 | 1.38E-4 | - 1.03E-4 | 4.39E-4 |
| MFDU | 3.159E-4 | 0.1011 | 2.20E-4 | 2.12E-4 | - 2.22E-4 | 6.13E-04 | 4.00E-5 | 1.91E-4 | - 3.40E-4 | 4.21E-4 |
| CountyBER | 0.2681 | 0.0046 | 0.2532 | 0.099 | 0.0686 | 0.4733 | - 0.4514 | 0.2228 | - 0.8486 | 0.1459 |
| CountyVAL | - 0.4646 | 0.0020 | - 0.4256 | 0.1675 | - 0.7553 | - 0.1075 | - 0.7375 | 0.6007 | - 1.992 | 0.4361 |
| | | | | | | | | | | |
| Intercept | 2.2889 | < 0.001 | 0.9874 | 0.1362 | 0.7385 | 1.255 | 2.184 | 0.1766 | 1.754 | 2.528 |
| | | | | | | | | | | |
| AIC | 9355 | | - | | | | - | | | |
| DIC | - | | 6637 | | | | 5788 | | | |

**Table 6: Parameter Estimates from Severe Crash Models**

| Variables | Non-Bayesian NB Model (Model-6a) | | Bayesian model without accounting spatial correlation (Model-6b) | | | | Bayesian model accounting spatial correlation (Model-6c) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Estimates | P-value | Mean | Std. Dev. | Bayesian Credible Interval | | Mean | Std. Dev. | Bayesian Credible Interval | |
| | | | | | 2.5% | 97.5% | | | 2.5% | 97.5% |
| SigInt | 0.1943 | < 0.001 | 0.2088 | 0.0311 | 0.1472 | 0.2704 | 0.1881 | 0.0268 | 0.1358 | 0.2407 |
| LnPop | 0.1015 | 0.0008 | 0.1062 | 0.0324 | 0.0389 | 0.1696 | 0.0545 | 0.0413 | - 0.0232 | 0.1335 |
| LnBasic | 0.0916 | 0.0003 | 0.0888 | 0.0277 | 0.0321 | 0.1425 | 0.0461 | 0.0259 | - 0.0049 | 0.0967 |
| LnRetail | 0.1421 | < 0.001 | 0.1827 | 0.0270 | 0.1288 | 0.2329 | 0.128 | 0.0253 | 0.0759 | 0.1776 |
| LnService | 0.1486 | < 0.001 | 0.1959 | 0.0293 | 0.1427 | 0.2569 | 0.1432 | 0.0282 | 0.0883 | 0.198 |
| LnMedInc | - 0.056 | < 0.001 | - 0.0416 | 0.0151 | -0.0682 | - 0.0088 | - 0.0155 | 0.0204 | - 0.0546 | 0.0275 |
| SFDU | - 1.885E-4 | 0.1831 | - 1.14E-4 | 1.55E-4 | - 4.20E-4 | 1.81E-4 | 1.05E-4 | 1.61E-4 | - 2.06E-4 | 4.03E-4 |
| MFDU | 3.639E4 | 0.0526 | 2.63E-4 | 2.17E-4 | - 1.66E-4 | 6.77E-4 | 9.38E-5 | 1.90E-4 | - 2.80E-4 | 4.69E-4 |
| CountyBER | 0.1451 | 0.1238 | 0.169 | 0.109 | - 0.0506 | 0.3824 | - 0.5593 | 0.3116 | - 1.134 | 0.0634 |
| CountyVAL | - 0.4466 | 0.0031 | - 0.4153 | 0.1697 | - 0.7523 | - 0.0776 | - 0.8207 | 0.5854 | - 1.986 | 0.278 |
| | | | | | | | | | | |
| Intercept | 1.3229 | < 0.001 | 0.3129 | 0.1366 | 0.0313 | 0.5721 | 1.385 | 0.2949 | 0.8422 | 1.968 |
| | | | | | | | | | | |
| AIC | 7273 | | - | | | | - | | | |
| DIC | - | | 5533 | | | | 1083 | | | |

**Table 7: Percent Increase of Crashes for Different Scenarios in Five-Year Interval**

| Alternate 1 Scenario | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Negative Binomial Model | 6.98 | 10.75 | 15.99 | 21.02 | 26.31 | 30.54 | | Negative Binomial Model | 5.67 | 9.41 | 14.34 | 19.07 | 24.14 | 28.24 |
| Spatial Bayesian Model | 2.15 | 5.34 | 9.55 | 13.53 | 17.27 | 20.36 | | Spatial Bayesian Model | 0.79 | 3.7 | 7.34 | 10.68 | 13.85 | 16.45 |
| Non-Spatial Bayesian Model | 4.14 | 9.55 | 16.6 | 23.54 | 30.21 | 35.88 | | Non-Spatial Bayesian Model | 2.41 | 6.75 | 12.61 | 18.4 | 24.05 | 28.84 |
| | | | | | | | | | | | | | | |
| **Alternate 2 Scenario** | | | | | | | | | | | | | | |
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Negative Binomial Model | 6.49 | 14.42 | 22.23 | 30.75 | 38.17 | 45.52 | | Negative Binomial Model | 5.17 | 12.39 | 19.68 | 27.93 | 35.12 | 42.41 |
| Spatial Bayesian Model | 2.05 | 7.34 | 13.19 | 18.12 | 22.34 | 26.26 | | Spatial Bayesian Model | 0.57 | 5.09 | 10.2 | 14.64 | 18.41 | 21.95 |
| Non-Spatial Bayesian Model | 4.35 | 14.92 | 28.51 | 40.67 | 51.18 | 62.14 | | Non-Spatial Bayesian Model | 1.79 | 9.98 | 20.91 | 31.08 | 39.79 | 48.95 |
| | | | | | | | | | | | | | | |
| **Alternate 3 Scenario** | | | | | | | | | | | | | | |
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Negative Binomial Model | 6.49 | 14.42 | 22.23 | 30.31 | 37.94 | 45.42 | | Negative Binomial Model | 5.17 | 12.39 | 19.68 | 27.5 | 34.96 | 42.48 |
| Spatial Bayesian Model | 2.05 | 7.34 | 13.19 | 18.08 | 22.52 | 26.45 | | Spatial Bayesian Model | 0.57 | 5.09 | 10.2 | 14.54 | 18.48 | 21.97 |
| Non-Spatial Bayesian Model | 4.35 | 14.92 | 28.51 | 40.04 | 51.14 | 61.41 | | Non-Spatial Bayesian Model | 1.79 | 9.98 | 20.91 | 30.39 | 39.59 | 48.07 |

**Table 8: Percent Increase of Crashes in Bernalillo and Sandoval Counties Based on Spatial Poisson-Lognormal Model Forecast (Predictions Calculated from the Region-Wide Model)**

| **Bernalillo County** | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Alternate 1 Scenario | 2.15 | 4.79 | 8.58 | 12.13 | 15.26 | 17.56 | | Alternate 1 Scenario | 0.89 | 3.33 | 6.65 | 9.69 | 12.37 | 14.41 |
| Alternate 2 Scenario | 2.51 | 7.79 | 13.32 | 17.66 | 21.4 | 24.9 | | Alternate 2 Scenario | 1.12 | 5.65 | 10.54 | 14.58 | 18.01 | 21.3 |
| Alternate 3 Scenario | 2.51 | 7.79 | 13.33 | 17.58 | 21.55 | 24.92 | | Alternate 3 Scenario | 1.12 | 5.65 | 10.54 | 14.41 | 18.03 | 21.16 |
| | | | | | | | | | | | | | | |
| **Sandoval County** | | | | | | | | | | | | | | |
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Alternate 1 Scenario | 8.73 | 18.9 | 27.96 | 36.42 | 45.63 | 55.17 | | Alternate 1 Scenario | 6.06 | 14.16 | 20.89 | 27 | 33.98 | 40.58 |
| Alternate 2 Scenario | 2.28 | 10.18 | 21.04 | 32.16 | 41.5 | 49.24 | | Alternate 2 Scenario | 0.81 | 7.41 | 15.51 | 23.66 | 30.52 | 36.1 |
| Alternate 3 Scenario | 2.28 | 10.18 | 21.04 | 32.95 | 41.86 | 50.45 | | Alternate 3 Scenario | 0.81 | 7.41 | 15.51 | 24.4 | 30.9 | 36.93 |

**Table 9: Crash Models Developed Specific for Bernalillo County**

| Variables | Bayesian model for Total Crashes accounting spatial correlation (Model-8a) | | | | Bayesian model for Severe Crashes accounting spatial correlation (Model-8b) | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std. Dev. | Bayesian Credible Interval | | Mean | Std. Dev. | Bayesian Credible Interval | |
| | | | 2.5% | 97.5% | | | 2.5% | 97.5% |
| SigInt | 0.2088 | 0.03 | 0.1494 | 0.2669 | 0.1966 | 0.0293 | 0.1387 | 0.2541 |
| LnPop | - 0.0161 | 0.0461 | - 0.1076 | 0.074 | 0.0086 | 0.0462 | - 0.0808 | 0.0992 |
| LnBasic | 0.0641 | 0.031 | 0.0035 | 0.125 | 0.0448 | 0.0306 | - 0.015 | 0.104 |
| LnRetail | 0.1269 | 0.0288 | 0.069 | 0.184 | 0.12 | 0.0277 | 0.0676 | 0.1754 |
| LnService | 0.1878 | 0.0306 | 0.1301 | 0.2505 | 0.1736 | 0.0298 | 0.1168 | 0.2333 |
| LnMedInc | 0.0216 | 0.0238 | - 0.0225 | 0.0682 | 0.0061 | 0.0248 | - 0.0422 | 0.0551 |
| SFDU | 3.14E-4 | 1.79E-4 | - 5.12E-5 | 6.56E-4 | 1.69E-4 | 1.84E-4 | - 1.94E-4 | 5.21E-4 |
| MFDU | 9.72E-5 | 1.96E-4 | - 2.83E-4 | 4.82E-4 | 7.13E-5 | 1.86E-4 | - 2.86E-4 | 4.45E-4 |
| | | | | | | | | |
| Intercept | 1.897 | 0.1835 | 1.502 | 2.244 | 0.9616 | 0.1738 | 0.6228 | 1.296 |

**Table 10: Crash Models Developed Specific for Sandoval County**

| Variables | Bayesian model for Total Crashes accounting spatial correlation (Model-9a) | | | | Bayesian model for Severe Crashes accounting spatial correlation (Model-9b) | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std. Dev. | Bayesian Credible Interval | | Mean | Std. Dev. | Bayesian Credible Interval | |
| | | | 2.5% | 97.5% | | | 2.5% | 97.5% |
| SigInt | 0.2092 | 0.1063 | 4.06E-4 | 0.4169 | 0.2203 | 0.111 | 0.0097 | 0.4356 |
| LnPop | 0.1503 | 0.0932 | - 0.0325 | 0.3266 | 0.1426 | 0.0867 | - 0.0316 | 0.3125 |
| LnBasic | 0.1497 | 0.1064 | - 0.0544 | 0.3565 | 0.1211 | 0.1082 | - 0.0861 | 0.3419 |
| LnRetail | 0.1349 | 0.0906 | - 0.0422 | 0.3183 | 0.1452 | 0.0914 | - 0.0325 | 0.3213 |
| LnService | 0.1297 | 0.0895 | - 0.0461 | 0.3014 | 0.0966 | 0.0903 | - 0.0801 | 0.2723 |
| LnMedInc | - 0.0415 | 0.0436 | - 0.1224 | 0.0525 | - 0.0571 | 0.039 | - 0.1302 | 0.0252 |
| SFDU | - 6.70E-5 | 4.07E-4 | - 8.60E-4 | 7.01E-4 | - 1.04E-4 | 3.94E-4 | - 8.59E-4 | 6.74E-4 |
| MFDU | 0.0018 | 0.0015 | - 0.001 | 0.0049 | 0.0017 | 0.0014 | - 9.89E-4 | 0.0045 |
| | | | | | | | | |
| Intercept | 1.13 | 0.3779 | 0.4106 | 1.875 | 0.5139 | 0.3739 | - 0.234 | 1.221 |

**Table 11: Correlation matrices for Crash Models Developed Specific for Bernalillo and Sandoval County**

**Bernalillo County - Total Crash Model (Model-8a)**

| | SigInt | LnPop | LnBasic | LnRetail | LnService | LnMedInc | SFDU | MFDU |
|---|---|---|---|---|---|---|---|---|
| SigInt | 1.000 | | | | | | | |
| LnPop | 0.033 | 1.000 | | | | | | |
| LnBasic | -0.090 | -0.075 | 1.000 | | | | | |
| LnRetail | 0.361 | -0.012 | -0.243 | 1.000 | | | | |
| LnService | 0.103 | -0.108 | 0.028 | 0.092 | 1.000 | | | |
| LnMedInc | -0.176 | **-0.813** | 0.133 | -0.061 | 0.081 | 1.000 | | |
| SFDU | -0.004 | **-0.625** | -0.013 | 0.036 | 0.134 | 0.356 | 1.000 | |
| MFDU | 0.109 | -0.372 | 0.063 | 0.088 | -0.001 | 0.139 | 0.305 | 1.000 |

**Bernalillo County - Severe Crash Model (Model-8b)**

| | SigInt | LnPop | LnBasic | LnRetail | LnService | LnMedInc | SFDU | MFDU |
|---|---|---|---|---|---|---|---|---|
| SigInt | 1.000 | | | | | | | |
| LnPop | 0.072 | 1.000 | | | | | | |
| LnBasic | -0.137 | -0.040 | 1.000 | | | | | |
| LnRetail | -0.027 | 0.044 | -0.290 | 1.000 | | | | |
| LnService | 0.013 | -0.110 | -0.077 | -0.248 | 1.000 | | | |
| LnMedInc | -0.086 | **-0.843** | 0.076 | -0.020 | 0.031 | 1.000 | | |
| SFDU | -0.043 | **-0.618** | -0.062 | -0.099 | 0.067 | 0.398 | 1.000 | |
| MFDU | -0.055 | -0.432 | -0.005 | -0.068 | 0.005 | 0.292 | 0.312 | 1.000 |

**Sandoval County - Total Crash Model (Model-9a)**

| | SigInt | LnPop | LnBasic | LnRetail | LnService | LnMedInc | SFDU | MFDU |
|---|---|---|---|---|---|---|---|---|
| SigInt | 1.000 | | | | | | | |
| LnPop | 0.103 | 1.000 | | | | | | |
| LnBasic | -0.041 | -0.004 | 1.000 | | | | | |
| LnRetail | -0.063 | 0.028 | -0.314 | 1.000 | | | | |
| LnService | -0.030 | -0.171 | -0.334 | -0.481 | 1.000 | | | |
| LnMedInc | -0.208 | **-0.616** | 0.029 | -0.096 | 0.118 | 1.000 | | |
| SFDU | -0.130 | **-0.629** | -0.173 | 0.107 | 0.020 | 0.282 | 1.000 | |
| MFDU | 0.032 | -0.091 | -0.081 | 0.050 | -0.008 | -0.080 | 0.090 | 1.000 |

**Sandoval County - Severe Crash Model (Model-9b)**

| | SigInt | LnPop | LnBasic | LnRetail | LnService | LnMedInc | SFDU | MFDU |
|---|---|---|---|---|---|---|---|---|
| SigInt | 1.000 | | | | | | | |
| LnPop | 0.131 | 1.000 | | | | | | |
| LnBasic | -0.056 | -0.029 | 1.000 | | | | | |
| LnRetail | -0.133 | 0.061 | -0.383 | 1.000 | | | | |
| LnService | -0.062 | -0.186 | -0.302 | -0.483 | 1.000 | | | |
| LnMedInc | -0.226 | **-0.617** | 0.075 | -0.094 | 0.148 | 1.000 | | |
| SFDU | -0.172 | **-0.634** | -0.112 | 0.025 | -0.004 | 0.278 | 1.000 | |
| MFDU | -0.232 | -0.122 | -0.035 | -0.040 | -0.022 | 0.010 | 0.103 | 1.000 |

**Table 12: Percent Increase of Crashes in Bernalillo and Sandoval Counties Based on Spatial Poisson-Lognormal Model Forecast (Predictions Calculated from Each County-Specific Model)**

| Bernalillo County | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Alternate 1 Scenario | 2.65 | 6.1 | 9.78 | 12.88 | 15.72 | 17.81 | | Alternate 1 Scenario | 1.5 | 4.19 | 7.44 | 10.34 | 12.9 | 14.78 |
| Alternate 2 Scenario | 4.32 | 7.94 | 12.38 | 16.4 | 19.76 | 23.1 | | Alternate 2 Scenario | 2.28 | 5.67 | 9.81 | 13.42 | 16.47 | 19.37 |
| Alternate 3 Scenario | 4.32 | 7.94 | 12.38 | 16.12 | 19.66 | 22.9 | | Alternate 3 Scenario | 2.28 | 5.67 | 9.81 | 13.22 | 16.42 | 19.27 |
| | | | | | | | | | | | | | |
| **Sandoval County** | | | | | | | | | | | | | |
| *Total Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 | | *Severe Crash Forecasts* | 2011-2015 | 2016-2020 | 2021-2025 | 2026-2030 | 2031-2035 | 2036-2040 |
| Alternate 1 Scenario | 10.34 | 21.86 | 31.18 | 39.27 | 48.35 | 58.15 | | Alternate 1 Scenario | 5.47 | 14.41 | 21.4 | 27.27 | 34.23 | 41.17 |
| Alternate 2 Scenario | 2.04 | 11.95 | 23.23 | 35.71 | 46.29 | 55.03 | | Alternate 2 Scenario | -1.05 | 7.02 | 15.56 | 24.59 | 32.23 | 38.51 |
| Alternate 3 Scenario | 2.04 | 11.95 | 23.23 | 36.77 | 46.55 | 56.01 | | Alternate 3 Scenario | -1.05 | 7.02 | 15.56 | 25.49 | 32.43 | 38.98 |

# FIGURES

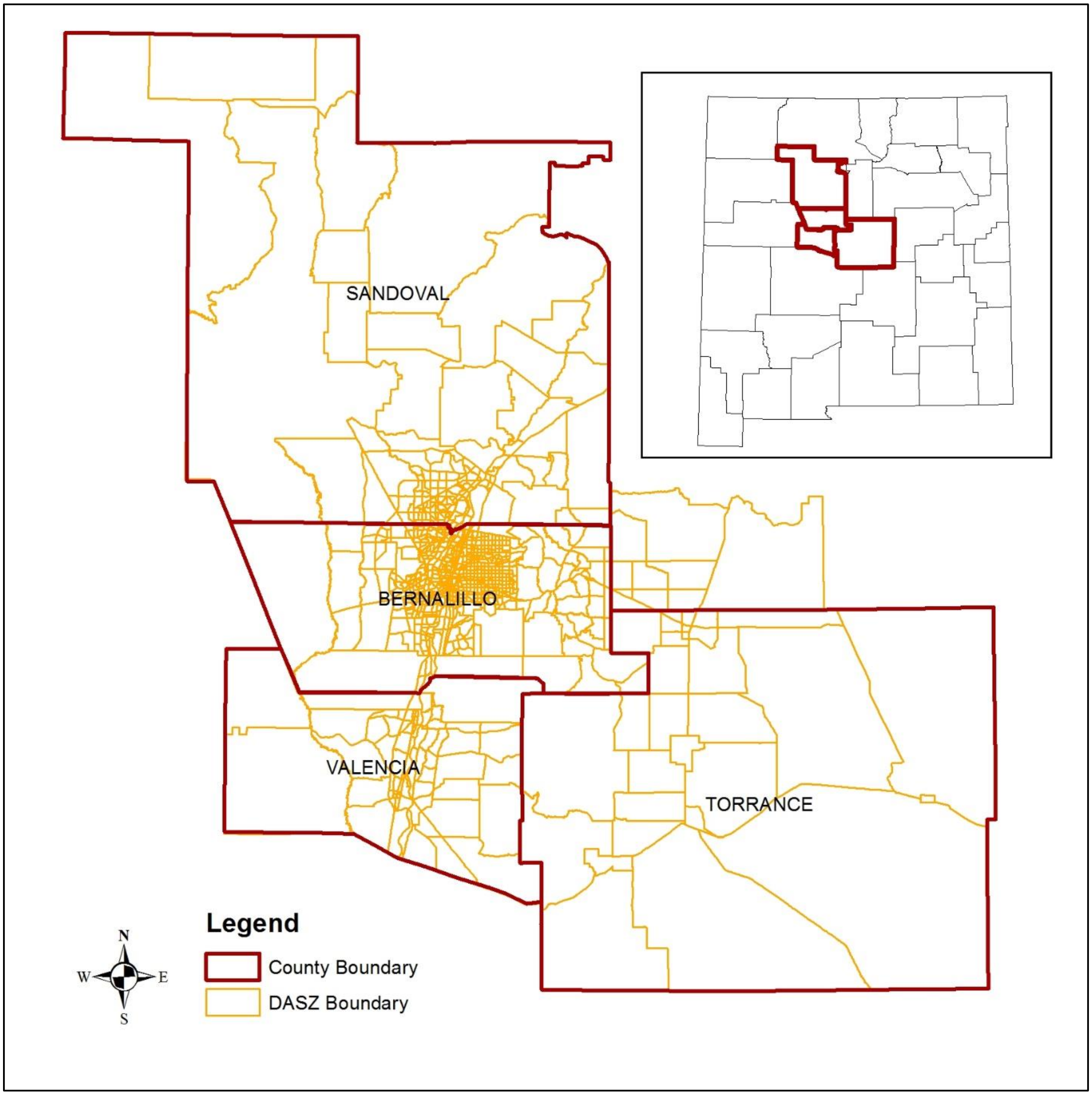


**Figure 1: Study Area**

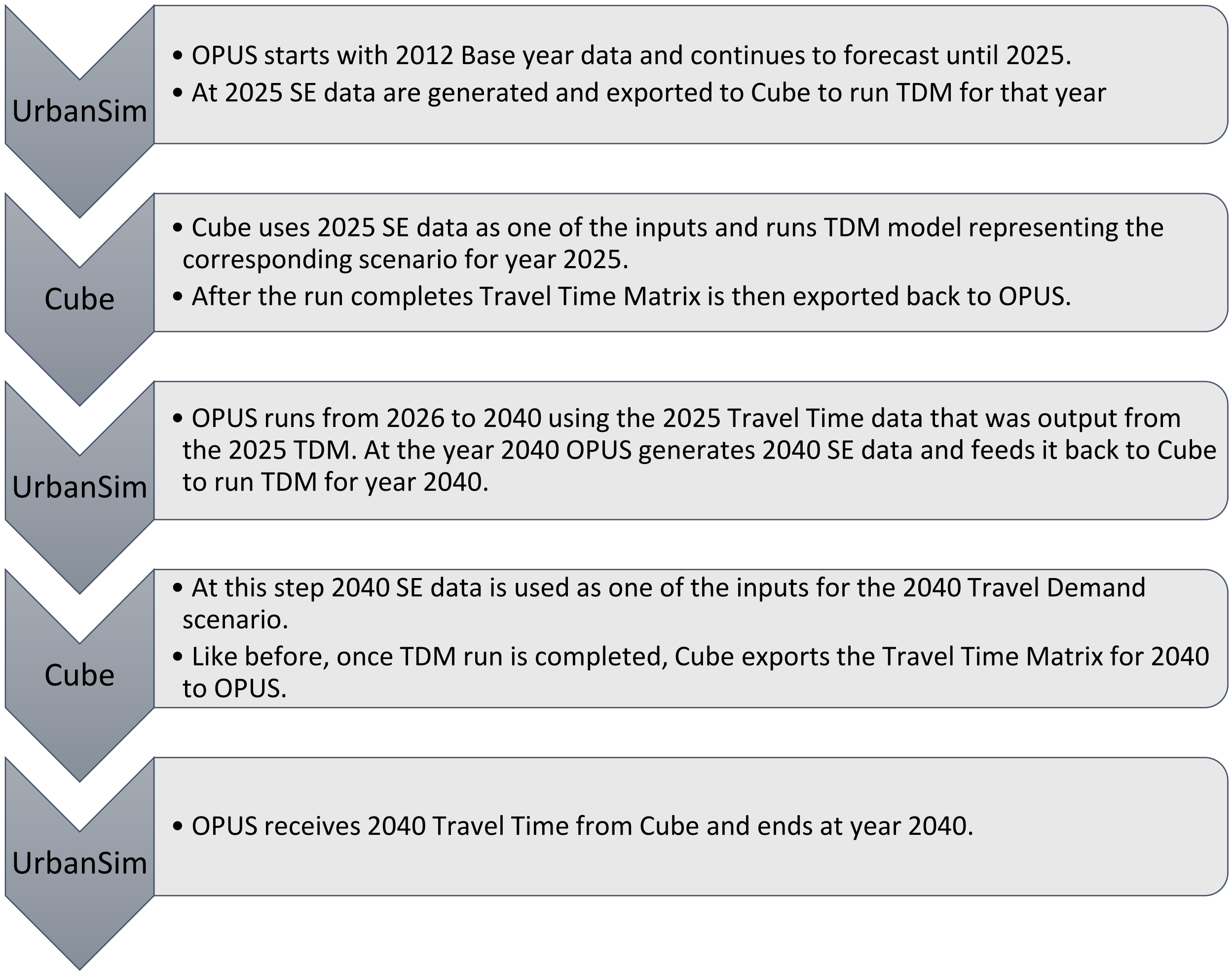


**Figure 2: UrbanSim-Cube interaction**