THE JOURNAL OF TRANSPORT AND LAND USE http://jtlu.org
VOL. 8 NO. 3 [2015] pp. 57–70 doi: 10.5198/jtlu.v8i3.614

JTLU

Assessing and integrating uncertainty into land-use forecasting

Hana Ševčíková[a]
Puget Sound Regional Council
University of Washington

Mark Simonson[b]
Puget Sound Regional Council

Michael Jensen[c]
Puget Sound Regional Council

Abstract: Uncertainty in land use and transportation modeling has received increasing attention in the past few years. However, methods for quantifying uncertainty in such models are usually developed in an academic environment and in most cases do not reach users of official forecasts, such as planners and policymakers. In this paper, we describe the practical application of a methodology called Bayesian melding and its integration into the land-use forecast published by the Puget Sound Regional Council, a metropolitan planning organization. The method allows practitioners to assess uncertainty about forecasted quantities, such as households, population, and jobs, for each geographic unit. Users are provided with probability intervals around forecasts, which add value to model validation, scenario comparison, and external review and comment procedures. Practical issues such as how many runs to use or assessing uncertainty for aggregated regions are also discussed.

Keywords: uncertainty; land-use forecast; Bayesian melding; UrbanSim; agent-based models; PSRC

1 Introduction

Uncertainty is an inherent aspect of any forecast—and for forecast users, a valuable one to be measured. Simulation models only approximate real systems. Input data or parameters are measured or estimated with varying degrees of accuracy, and models (and the systems they represent) are often explicitly stochastic. However, official forecasts are usually published as point predictions, leaving end users to make their own assumptions about forecast reliability. Both forecasters and end users would be better served by comprehensible metrics that illustrate the limits of the forecast.

For greatest practical value, uncertainty metrics should pertain to the units being forecasted. Forecasters have typically quantified uncertainties only in some input components, if at all. Goodness-of-fit statistics for models within a system and margins of error for inputs and parameters have diagnostic value to the modeler, but they do little to explain the likelihood of alternative future outcomes. "High" and "low" scenarios from varying inputs give the appearance of bounds but have no statistical basis. Even simply repeating a simulation multiple times does not result in an accurate assessment of uncertainty (Ševčíková et al. 2007).

Several disciplines have made strides in assessing and communicating forecast uncertainty, but practical use outside of academia remains rare. Statistics Netherlands' probabilistic population projections (de Beer and Alders 1999) have been among the first. Recently, the United Nations Population

[a] HSevcikova@psrc.org
[b] MSimonson@psrc.org
[c] MJensen@psrc.org

Division also included probabilistic projections in their official forecasts (Gerland et al. 2014; Raftery et al. 2012; United Nations 2013). Pioneering academic studies have been written on the topic in the context of land use and transportation forecasting. Examples include Brown et al. (2005); Duthie et al. (2010); Pontius Jr and Spencer (2005); Pradhan and Kockelman (2002); Rodier (2005); Rodier and Johnston (2002); Tayman (1996, 2011); Zhao and Kockelman (2002). However, until now such metrics have not reached the point of inclusion in forecasts.

Our article describes what we believe to be the first incorporation of uncertainty assessment into the development of an official land-use forecast published by a metropolitan planning organization (MPO). The methodology for assessing uncertainty in the forecast is based on work by Ševčíková et al. (2007) and Ševčíková et al. (2011) and is implemented as open source python code. The resulting uncertainty metrics played an important role in the process of producing the forecast, both during model validation and during a subsequent external review and comment period. They have been favorably received by forecast users, and we expect such information will play an increasing role in the interpretation of future forecasts and scenario simulations.

The paper is organized as follows: Section 2 provides the forecasting context at Puget Sound Regional Council (PSRC), including an overview of UrbanSim (Waddell 2002; Waddell et al. 2003), the land-use model system involved. Section 3 describes the methodology. The application of the method and its results, including aspects such as the impact of random numbers, and aggregations are discussed in Section 4. Section 5 summarizes the implementation. Ways the forecast uncertainty were used are discussed in Section 6, and Section 7 draws initial conclusions and outlines future efforts.

2 Land-use forecasting at PSRC

The immediate context for this work is land-use modeling at PSRC using the UrbanSim urban simulation system. UrbanSim is operational in several urban areas in the United States (Waddell 2010, 2011; Waddell et al. 2007) and Europe (Felsenstein et al. 2010). The system is implemented as a set of interacting agent-based models that represent the major actors and choices in the urban system, including households choosing residential locations, business choices of employment location, worker choices of jobs, and developer choices of locations and types of real estate development. The model system microsimulates the annual evolution in locations of individual households and jobs, including the connection between them, and the evolution of the real estate within each individual geography as the result of actions by real estate developers.

UrbanSim is implemented as part of the OPUS framework, or Open Platform for Urban Simulation (Waddell et al. 2005), in Python code. Its flexible construction and modularity allows it to be easily adapted to regional conditions; such adaptations can be implemented as separate Python packages. OPUS also allows connections between the land-use models and an external travel model (Waddell et al. 2010).

PSRC's implementation uses exogenous household and job control totals (from a macroeconomic model) and locates them into existing or simulated new/redeveloped built space within the region. Operating on a parcel level, the models recognize detailed development constraints, attributes of existing buildings, job counts by sector, and synthesized attributes of households and persons, among other inputs. Development in UrbanSim is influenced by price, which is altered throughout the simulation using hedonic regression (Rosen 1974). Multinominal logit (MNL) models (McFadden 1974, 1978, 1981) are used to model discrete location choice of households and of jobs and to link workers to jobs (Wang et al. 2011). The models are summarized in Table 1. Model segmentation or sampling strategies can be adjusted appropriately to the phenomenon at hand. Maximum likelihood estimation is used using primarily 2006 observed data, and each model cycle simulates a single year. For the 2013 land-use baseline forecast mentioned in this paper, PSRC used a 2000 base year, periodic travel

Table 1: UrbanSim models in the PSRC application.

Model	Method	Description
Real estate price	Hedonic regression	Predicts prices of parcels.
Expected sale price	Hedonic regression	Predicts prices of possible real estate proposals.
Development proposal choice	Weighted random sampling	Chooses real estate proposals to be built (including redevelopment) based on the expected sales prices of proposals.
Building construction	Rule based	Demolishes buildings (for redevelopment) and builds new buildings according to the chosen proposals.
Household transition	Random sampling	Creates and removes households. It is driven by macroeconomic predictions, which provide regional control totals for households grouped by their characteristics.
Employment transition	Random sampling	Creates and removes jobs. It is driven by macroeconomic predictions, which provide regional control totals for jobs grouped by employment sectors.
Household relocation	Weighted random sampling	Determines households for moving.
Household location choice	Multinomial logit with random sampling of alternatives	Locates moving households into buildings.
Employment relocation	Weighted random sampling	Determines jobs for moving.
Employment location choice	Multinomial logit with random sampling of alternatives	Locates moving jobs into buildings.
Work at home choice	Binary logit	Simulates workers decision to work at home or out of home.
Workplace relocation	Rule based	Simulates workers decision to change job.
Workplace choice	Multinomial logit with random sampling of alternatives	Assigns workers to jobs.

model integration to update accessibility variables, and a 2040 horizon year. This is PSRC's first published forecast using UrbanSim; prior forecasts used adjusted output from DRAM/EMPAL and did not include uncertainty metrics.

In addition to internal use, PSRC's published forecast products are widely used throughout the region for purposes including travel demand forecasting, utility forecasting, and capital improvement planning. PSRC's 6300-square-mile forecasting area is currently home to 3.78 million people; its constituent jurisdictions include four counties, 72 cities, two tribes, and four port districts (a map of the region can be seen in Figure 1). The agency serves as the federally designated MPO for this area, and it also has statutory responsibilities for growth management and economic development. The forecast was developed with input from jurisdictional staff serving on a technical advisory committee, and it underwent two extensive periods of public review (the first focused on inputs and the second focused on forecast results) prior to publication.

3 Methodology

To assess uncertainty in the forecast, we use Bayesian melding, a methodology initially developed to calibrate uncertainty in deterministic model systems by Raftery et al. (1992, 1995) and Poole and Raftery (2000). Ševčíková et al. (2007) and Ševčíková et al. (2011) adapted the method to stochastic models in a land-use and transportation forecasting context. The approach in this paper closely follows the above publications and is summarized in this section.

Figure 2 may help to communicate the basic concept developed for deterministic models. There is a prior distribution of model inputs $q(\Theta)$ from which we draw input values Θ_i for $i = 1, \ldots, I$. The model runs I times from the starting point to the present and for each input Θ_i it produces as output the quantity of interest, Φ_i. The model can be viewed as a mapping, M, from the space of inputs to the space of outputs, which we denote by $\Phi = M_\Phi(\Theta)$. The "present" time is defined as

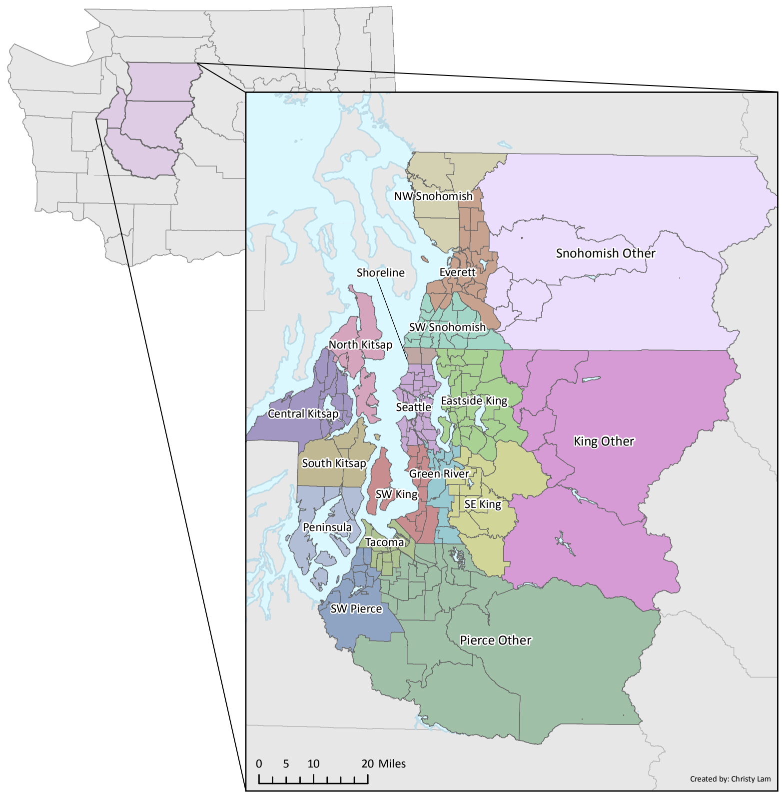


Figure 1: The PSRC region. Black lines delimit 219 forecast analysis zones, which are color-grouped into 18 large areas.

a time point for which we have observed data available. We use the observed data, denoted by y, to compute a weight w_i for each input Θ_i: $w_i = L(\Phi_i)$. Here, $L(\Phi_i)$ is the likelihood of the model outputs given the observed data, $L(\Phi_i) = \text{Prob}(y|\Phi_i)$. For each of the I runs, the model is run forward until a future time for which we make a prediction. The results of the ith model run are denoted by Ψ_i. The posterior distribution of Ψ is approximated by a discrete distribution with values Ψ_i having probabilities proportional to w_i.

In stochastic models, the conditional distribution of the model outputs, Φ, given the model inputs, Θ (which is a point mass at $M_\Phi(\Theta)$ for deterministic models), becomes a probability distribution. This distribution incorporates both the stochastic nature of the model outputs and the variability of the model error.

Let k denote an index of a geographic unit for which observed data y is available, i indexes the simulation run, and l indexes the quantity of interest. Then the model is defined as:

$$(y_{kl}|\Theta = \Theta_i) = \mu_{ikl} + a_l + \epsilon_{ikl}, \text{ where } \epsilon_{ikl} \overset{iid}{\sim} N(0, \sigma^2_{il}), \tag{1}$$

for $i = 1, \ldots, I$, $k = 1, \ldots, K$ and $l = 1, \ldots, L$. The quantity μ_{ikl} is the expected value of y_{kl} under the model given Θ_i, ϵ_{ikl} denotes the model error, and a_l is the overall bias in the model predictions

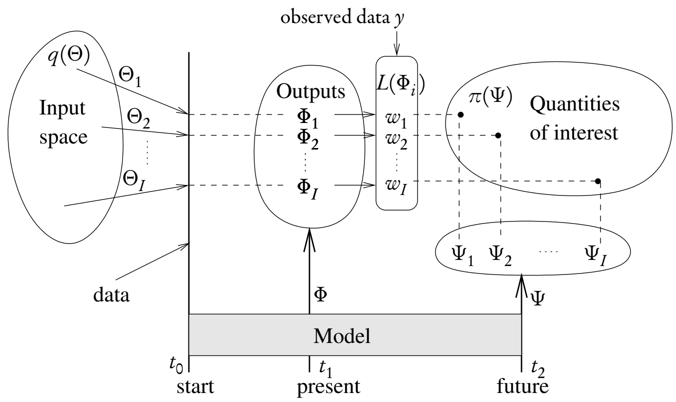


Figure 2: Illustration of the Bayesian melding method for deterministic models. The uncertain model inputs, Θ, refer to the starting time of the simulation, t_0, and the outputs, Φ. The data relevant to the outputs, y, are observed at the "present" time, t_1, while the quantities of interest, Ψ, refer to the future, t_2. The quantities Θ_i, Φ_i, and Ψ_i refer to the ith simulated values of inputs, outputs, and quantities of interest, respectively.

of the lth output. The variance σ^2_{il} and bias a_l are estimated by their sample equivalents:

$$\hat{\sigma}^2_{il} = \frac{1}{K}\sum_k (y_{kl} - \hat{a}_l - \hat{\mu}_{ikl})^2, \text{ and} \tag{2}$$

$$\hat{a}_l = \frac{1}{IK}\sum_{i,l}(y_{kl} - \hat{\mu}_{ikl}), \tag{3}$$

where $\hat{\mu}_{ikl}$ is the predicted value of y_{kl} from the ith simulation run.

This yields a conditional predictive distribution of our quantity of interest:

$$y_{kl}|\Theta_i \sim N(\hat{a}_l + \hat{\mu}_{ikl}, \hat{\sigma}^2_{il}) \tag{4}$$

with the weights computed as

$$w_i \propto p(y|\Theta_i) = \prod_{l=1}^{L}\prod_{k=1}^{K}\frac{1}{\sqrt{2\pi\hat{\sigma}^2_{il}}}\exp\left[-\frac{1/2(y_{kl} - \hat{a}_l - \hat{\mu}_{ikl})^2}{\hat{\sigma}^2_{il}}\right]. \tag{5}$$

The quantities $\hat{\sigma}^2_{il}$, $\hat{a}_l$, and w_i in Equations 2, 3, and 5 are estimated at the "present" time t_1.

The marginal distribution of the lth quantity of interest, Ψ_{kl}, in the future year t_2, is given by a mixture of normal distributions:

$$\pi(\Psi_{kl}) = \sum_{i=1}^{I} w_i N\left(f_a(\hat{a}_l, b^a_l) + \Psi_{ikl}, f_v(\hat{\sigma}^2_{il}, b^v_l)\right), \quad k = 1,\ldots,K,\ l = 1,\ldots,L. \tag{6}$$

Here, b^a_l and b^v_l denote propagation factors of the bias and the variance, respectively, over the time period $[t_1, t_2]$ for indicator l. Furthermore, $f_a, f_v \in \{*, +\}$ are functions determining how the propagation is applied.

To obtain uncertainty on a higher-level geography, one can either directly derive the posterior distribution using aggregated simulation results and calibrate on aggregated observed data. Or, one can draw a value for each k from the posterior distribution in Equation 6 and aggregate to the desired geography. Repeating such draws many times approximates the posterior distribution of the quantity of interest on aggregated geography.

4 Application and results

4.1 Data and settings

4.1.1 Data

Currently, PSRC's standard UrbanSim simulation begins in year 2000 (base year, t_0). UrbanSim requires highly disaggregate base year data. As previously noted, the Bayesian melding method also requires observed data (y) at the "present" time, t_1, or recent validation year, which can be less detailed. We used household, employment, and population counts aggregated to 938 traffic analysis zones (TAZ) for the 2010 validation year. These were aggregated further into 219 forecast analysis zones (FAZ) and 18 FAZ large areas (see map in Figure 1), the geographic units at which the forecasts are currently published.

4.1.2 Time

Starting from $t_0 = 2000$, we will use $t_1 = 2010$ as the present year (in which Equations 2, 3, and 5 are evaluated) and run the simulation forward until $t_2 = 2040$ (for which Equation 6 is applied). Furthermore, we use 2006 and 2008 as additional values for t_1 in order to estimate the propagation factors b_l^a and b_l^v.

4.1.3 Number of runs

One of the conclusions in Ševčíková et al. (2007) is that variation in model inputs or the random seed accounts for little of the total uncertainty. That suggests the parameter I can be relatively small. Given the long run-time of PSRC's (land use/travel model) integrated simulation, we reduced the number of runs to $I = 1$. The resultant loss in precision is small, as will be shown in Section 4.2. This also eliminates the need to compute the weights (Equation 5). The posterior distribution $\pi(\Psi_{kl})$ in Equation 6 then consists of only one (normal) component. For simplicity, in subsequent text we will eliminate the index i where appropriate.

4.1.4 Quantities of interest

The method accounts for uncertainty in forecast quantities for which there is validation data—i.e., indicators measured for a more recent year than the base year (i.e., present). In the land-use forecast, we used three indicators, namely number of households ($l = 1$), population ($l = 2$), and number of jobs ($l = 3$). For brevity, we omit population from the display of results, due to its similarity with the household results.

4.1.5 Transformation

Previous research in Ševčíková et al. (2007) showed that for quantities that represent counts, as it is the case of our chosen indicators, a square root transformation should be used in order to obtain approximately constant variance of the model errors. In practice, this means that quantities y, $\hat{\mu}$, and Ψ in Equations 2, 3 and 6 are all converted using a square root transformation. For example, using $l = 3$ and $k = 1$, $\hat{\mu}_{kl}$ is the square rooted number of jobs in zone 1 simulated at time t_1, and Ψ_{kl} is the square rooted number of jobs in zone 1 simulated at time t_2.

4.1.6 Calibration geography

Our desired geography for publishing forecast uncertainty is FAZ as well as FAZ large area. However, we have observed data on more disaggregated geography, namely TAZ. We will compare results from the two types of aggregations mentioned in Section 3 (i.e., a direct calibration on FAZ and FAZ large area levels with results simulated from distributions obtained on the TAZ level).

4.2 Results

4.2.1 Calibration and propagation factors

Applying Equations 2 and 3 at present time $t_1 = 2010$ for the three quantities of interest on the TAZ and FAZ geography yield results shown in Table 2.

Table 2: Results of the calibration step at time t_1. Here, Equations 2 and 3 were applied to the indicators number of households ($l = 1$), population ($l = 2$), and number of jobs ($l = 3$), all on the square root scale.

	$\hat{\sigma}^2_{l=1}$	$\hat{\sigma}^2_{l=2}$	$\hat{\sigma}^2_{l=3}$	$\hat{a}_{l=1}$	$\hat{a}_{l=2}$	$\hat{a}_{l=3}$
TAZ	7.78	37.11	52.81	−0.11	0.43	0.27
FAZ	13.85	39.39	46.19	0.18	−0.30	0.89

We used the observed data at $t_1 = 2006$ and $t_1 = 2008$ to compute the same quantities and determine the amount of propagation per year. For b^v, the propagation factor of the variance, we used multiple available simulation runs that differed in input parameters, data, or model structure. We then took the median value of the resulting propagations. This yield values of $b_1^v = 0.4d$, $b_2^v = 3.9d$ and $b_3^v = 2.2d$ where $d = t_2 - t_1 = 30$. These values are to be added to the variance—i.e., f_v in Equation 6 corresponds to the "+" operator.

For b^a, the propagation factor of the bias, we have not observed any systematic propagation. Furthermore, as mentioned in Section 2, the total number of households and jobs in the region is controlled by regional control totals obtained from macroeconomic forecasts. Adding a bias to the predictions would consistently increase or decrease the mean outcomes; thus, the sum of the means across the region would be inconsistent with the control totals. Therefore, we omit the bias and condition our predictions on the control totals. Our experiments show that applying the bias from Table 2 makes little difference in the predictions.

4.2.2 Varying random seed

To investigate to what extent our results change when we vary the random number sequence, we ran multiple runs with different random seeds and used the methodology from Section 3 with $I > 1$ to derive probability intervals in 2040. Figure 3 compares 95 percent intervals of such set of runs (red lines) with ones derived from only one run—i.e., when $I = 1$ (blue lines) for selected FAZs. Empty circles mark values from the multiple runs. Dots within the intervals mark posterior median of the distributions. For each indicator, we selected 20 zones with the largest difference between the results from multiple runs. That is, if only one run is used, the first zone from the left could experience the largest difference when switching to a different random seed, and all remaining (199 unplotted) zones will have a difference smaller than that of the zone on the right margin of the graph.

As can be seen, adding additional runs that differ in random seed does not impact the width of the intervals in most cases. This confirms prior research, which found random seed variation is not a major source of uncertainty—most of the uncertainty Bayesian melding is accounting for comes from other sources. However, the distribution becomes wider for zones with larger differences, as can be seen in the figure for the first few zones from the left, especially in the case of jobs, which suggests that repeating the simulation a few times might be advantageous to capture such behavior. However, as stated previously, for simplicity, in our published forecast we use $I = 1$, in which case the intervals are centered around UrbanSim outputs. The figure shows that such intervals contain the multiple runs results in vast majority of cases.

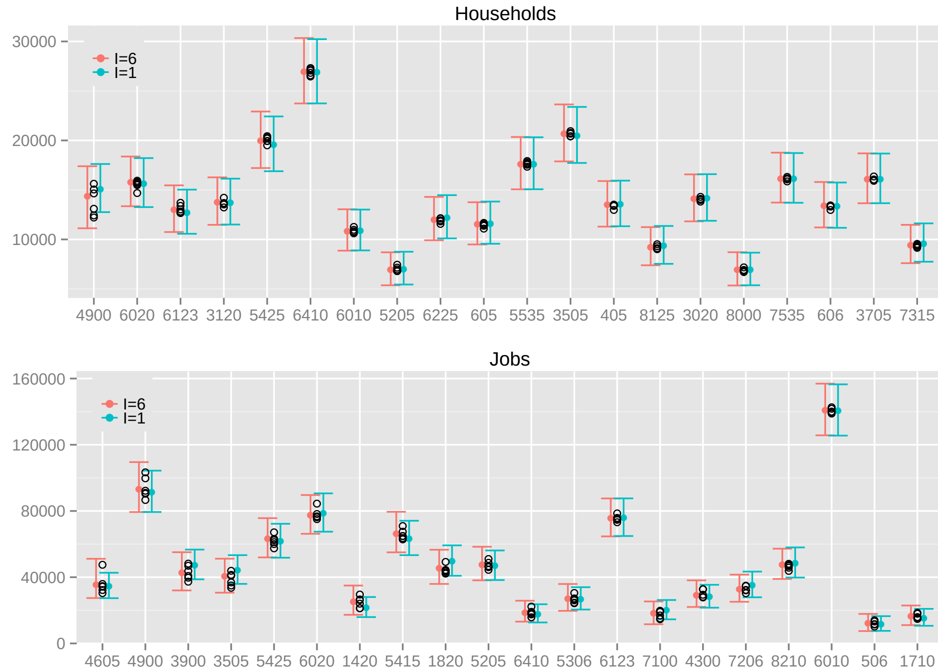


Figure 3: Ninety-five percent probability intervals in 2040 for a simulation with $I = 6$ (red bars) and $I = 1$ (blue bars), with medians marked by dots, for 20 FAZs with the largest difference between results from multiple runs (empty circles). The multiple runs varied in random seed.

4.2.3 Aggregation

As mentioned in Section 3, there are two ways of deriving probability intervals for an aggregated geography:

1. *By calibration*: Observed counts and simulated results are aggregated to the desired geography and the methodology from Section 3 is then applied.

2. *By simulation*: The methodology is applied to the disaggregated geography, for example TAZ. Then for each such zone k and indicator l, a large number, say J, of random numbers is sampled from the distribution in Equation 6—i.e., in our case $N(\Psi_{kl}, \hat{\sigma}_l^2 + b_l^v d)$ with $d = 30$ (centered around the [square root] transformed simulation result using the estimated variance propagated into 2040). These numbers are then transformed back to the original scale. For each l, the resulting $K \times J$ matrix is aggregated along the first (spacial) dimension to the desired geography and probability intervals can be derived from the second dimension. Optionally, a scaling of the disaggregated matrix can be done to meet the regional control totals.

Our experiments show that the probability intervals derived by simulation are somewhat smaller than when the calibration method is used. This suggests that the calibration method accounts for additional uncertainty that is due to correlation between zones within their aggregates, and thus is preferable. Figure 4 shows the difference for two indicators aggregated to FAZ large areas for the target year 2040. While the calibration method uses observed data on the FAZ large area level directly, the simulation method uses the TAZ results to simulate and aggregate to the FAZ large area level.

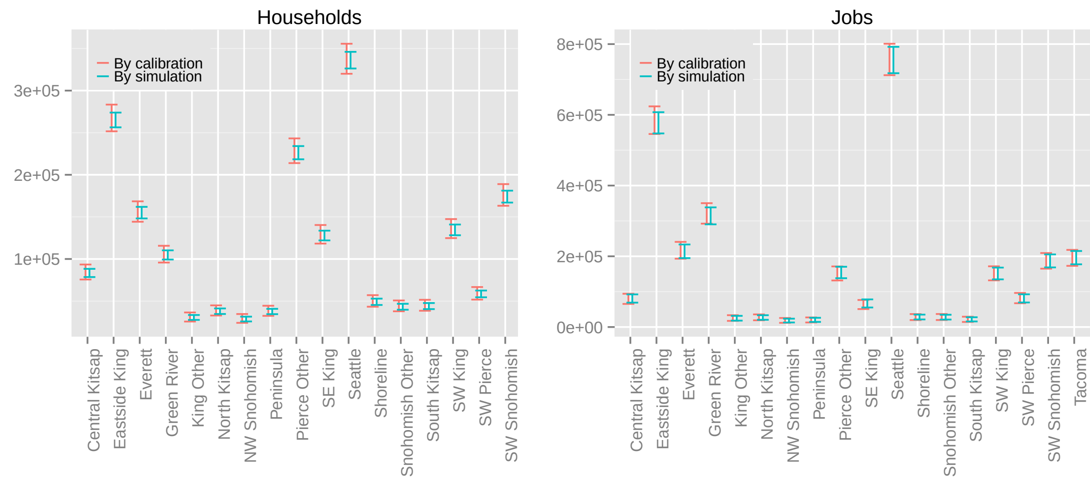


Figure 4: Ninety-five percent probability intervals for the two methods of aggregation: by calibration (red) and by simulation (blue). Results from the target prediction year (2040) are aggregated to the 18 FAZ large areas. The *by simulation* intervals are derived by aggregating up from the TAZ geography.

5 Implementation

Opus framework is available from the Subversion repository https://svn.urbansim.org. Most of the generic functionality resides in the package **opus_core**. Within this package, the Bayesian melding methodology is implemented in the module `bayesian_melding.py`, with the main class called `BayesianMelding`. The main method, called `compute_weights`, estimates the variance (Equation 2), bias (Equation 3) and weights (Equation 5). These parameters can be exported into a file using the method `export_bm_parameters` and re-used in a later session using the class `BayesianMeldingFromFile`. Other convenience methods are available—for example, for generating posterior distribution for future time point (Equation 6) or exporting probability intervals. The module contains a unit test called `test_bayesian_melding`, which shows an example of how to use these classes and their methods.

6 Using uncertainty

With information about how confident we can be in our forecast results, new ways become available to evaluate forecasts—and particularly, to compare multiple forecasts. Here we describe some ways in which we integrated uncertainty analysis into our land-use forecasting process.

6.1 Model validation

Forecasting production runs are usually preceded by a tedious phase of model estimation and testing, in which various alternatives of model structure as well as inputs and specifications are tested. In a complex system of models such as UrbanSim, change to a single component can have significant effects on other components, so it becomes a non-trivial task to evaluate the effect of changes on the simulation outcomes.

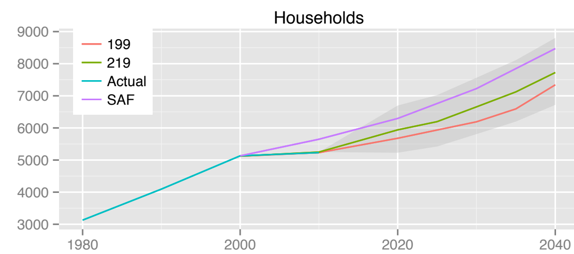
Our approach offers a straightforward way to compare two or more runs, using a single measure per run—namely, the weights w_i from Equation 5. The index i enumerates here the various runs to be compared. Intuitively, the higher the weight, the better the run validates to the observed data at t_1. The advantage of this approach is that it incorporates information about multiple quantities of interest into one measure. To investigate the impacts relative to each quantity of interest l, one can compare the variances $\hat{\sigma}^2_{il}$ from Equation 2 for each i. The larger the variance, the larger the distance of simulated results to the observed data.

6.2 Output review and refinement

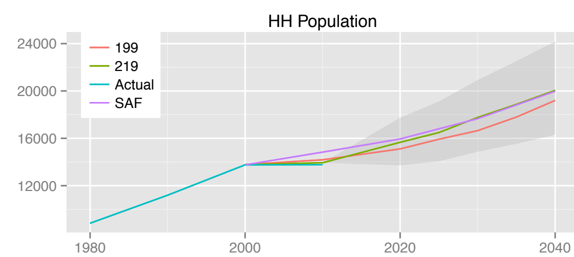
To facilitate forecast review, PSRC circulated plots that featured for each FAZ the forecast result with derived uncertainty in the form of 80 percent probability intervals, alongside past trends and previous forecasts (screenshot in Figure 5). The technical advisory committee provided us with comments in the form of values that are to be expected in these geographies in the future according to the experts. Comments that fell within the 80 percent probability interval were more likely to be accepted

6506 – North Bend area / Snoqualmie area

	Actual	SAF	run 199	run 219	Refined	low 219	high 219
1980	3,129						
1990	4,095						
2000	5,131	5,131	5,122	5,122			
2010	5,231	5,648	5,231	5,248			
2020		6,295	5,678	5,940		5,228	6,698
2030		7,225	6,191	6,656		5,807	7,563
2040		8,470	7,342	7,725		6,717	8,803



	Actual	SAF	run 199	run 219	Refined	low 219	high 219
1980	8,826						
1990	11,193						
2000	13,758	13,758	13,784	13,784			
2010	13,767	14,832	14,185	13,936			
2020		15,944	15,104	15,662		13,714	17,740
2030		17,657	16,647	17,761		14,859	20,922
2040		19,976	19,198	20,050		16,299	24,189



	Actual	SAF	run 199	run 219	Refined	low 219	high 219
1980	2,482						
1990	2,587						
2000	3,713	3,520	3,713	3,713			
2010	5,051	4,281	5,033	5,051			
2020		5,177	6,286	5,865	7,121	4,607	7,275
2030		5,930	7,912	7,401	9,334	5,692	9,334
2040		6,902	11,971	12,490		9,918	15,358

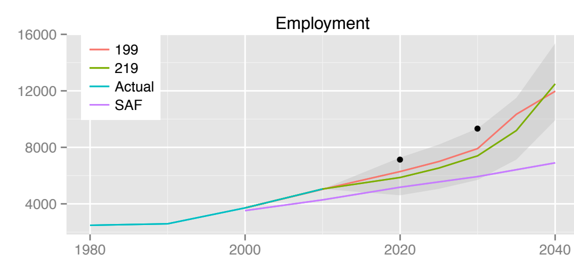


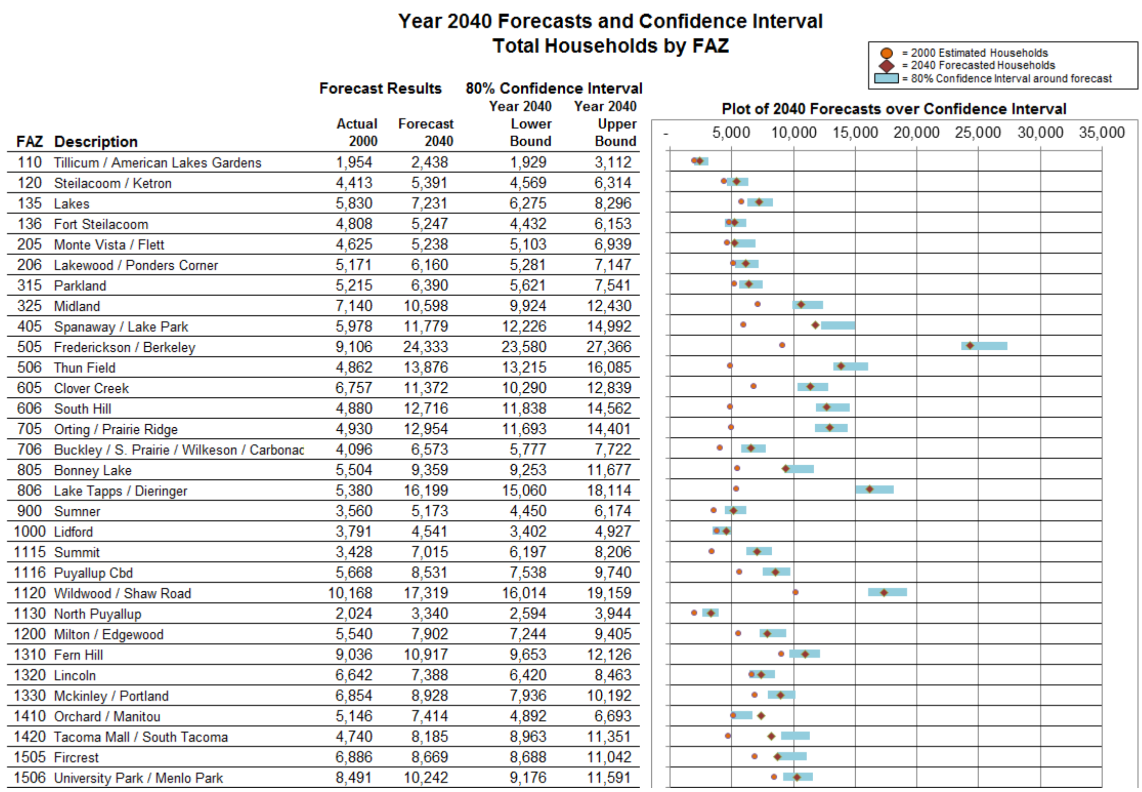
Figure 5: Screenshot of a time series report with indicators and probability intervals displayed as a function of time for a selected FAZ. The green lines (run 219) show the current forecast median with the 80 percent probability intervals as the shaded area. The red (run 199) and purple lines (SAF) show the previous forecasts, and the blue line represents the trend data. Manual adjustments are marked by black dots (denoted as "refined" in the table).

as adjustments prior to the final release (adjustments marked in Figure 5 using black dots). Comments outside the interval were understood to suggest errors in model inputs or structure, and thus required additional investigation and corroboration.

Note in the final reporting, the uncertainty progression starts at 2010 (t_1) as opposed to 2000 (t_0). In our application, a so-called reset step is added, in which we set the spacial distribution of the land-use actors at t_1 to observed data. Thus, we use simulation results from t_1 to obtain all quantities needed by the methodology, reset the t_1 data to observed data, and simulate to t_2. Then, the propagation factor discussed in Section 4.2 is reduced by $t_1 - t_0 = 10$—i.e., $d = t_2 - t_1 - 10 = 20$.

6.3 Forecast publication

The final derived probability intervals were incorporated into a land-use forecast validation report (Puget Sound Regional Council 2013b), released concurrent with the full forecast products package (Puget Sound Regional Council 2013a). Figure 6 shows a screenshot of the published information. It includes 80 percent probability intervals (as lower and upper bounds) around the simulated values. The choice of 80 percent bounds is due to its straight-forward interpretability: there is one chance in ten that the true value falls below the interval and one chance in ten that it falls above the interval.



Year 2040 Forecasts and Confidence Interval
Total Households by FAZ

FAZ	Description	Forecast Results: Actual 2000	Forecast Results: Forecast 2040	80% Confidence Interval: Year 2040 Lower Bound	80% Confidence Interval: Year 2040 Upper Bound
110	Tillicum / American Lakes Gardens	1,954	2,438	1,929	3,112
120	Steilacoom / Ketron	4,413	5,391	4,569	6,314
135	Lakes	5,830	7,231	6,275	8,296
136	Fort Steilacoom	4,808	5,247	4,432	6,153
205	Monte Vista / Flett	4,625	5,238	5,103	6,939
206	Lakewood / Ponders Corner	5,171	6,160	5,281	7,147
315	Parkland	5,215	6,390	5,621	7,541
325	Midland	7,140	10,598	9,924	12,430
405	Spanaway / Lake Park	5,978	11,779	12,226	14,992
505	Frederickson / Berkeley	9,106	24,333	23,580	27,366
506	Thun Field	4,862	13,876	13,215	16,085
605	Clover Creek	6,757	11,372	10,290	12,839
606	South Hill	4,880	12,716	11,838	14,562
705	Orting / Prairie Ridge	4,930	12,954	11,693	14,401
706	Buckley / S. Prairie / Wilkeson / Carbonac	4,096	6,573	5,777	7,722
805	Bonney Lake	5,504	9,359	9,253	11,677
806	Lake Tapps / Dieringer	5,380	16,199	15,060	18,114
900	Sumner	3,560	5,173	4,450	6,174
1000	Lidford	3,791	4,541	3,402	4,927
1115	Summit	3,428	7,015	6,197	8,206
1116	Puyallup Cbd	5,668	8,531	7,538	9,740
1120	Wildwood / Shaw Road	10,168	17,319	16,014	19,159
1130	North Puyallup	2,024	3,340	2,594	3,944
1200	Milton / Edgewood	5,540	7,902	7,244	9,405
1310	Fern Hill	9,036	10,917	9,653	12,126
1320	Lincoln	6,642	7,388	6,420	8,463
1330	Mckinley / Portland	6,854	8,928	7,936	10,192
1410	Orchard / Manitou	5,146	7,414	4,892	6,693
1420	Tacoma Mall / South Tacoma	4,740	8,185	8,963	11,351
1505	Fircrest	6,886	8,669	8,688	11,042
1506	University Park / Menlo Park	8,491	10,242	9,176	11,591

Figure 6: Screenshot of the published forecast that includes 80 percent probability intervals.

7 Discussion

Inclusion of uncertainty measures derived from a statistically grounded approach contributed to the success of the PSRC 2013 Land Use Baseline Forecast. Their value has extended beyond the purposes of review or validation; they remain salient to interpretation of the eventual forecast product as well.

Since land-use models are not used in isolation, concomitant uses suggest themselves, particularly in terms of travel modeling. Several publications such as Duthie et al. (2010) cite land-use uncertainty as a major source of travel model forecast uncertainty. Quantifying this via the approach described in this paper may help toward propagating the uncertainty into travel model outcomes, as was done by Ševčíková et al. (2011), for example.

The way simulation models are regarded and used is undergoing a gradual change, one which the inclusion of uncertainty metrics may accelerate and deepen. Behavioral simulation models like UrbanSim may, as a result of their sophistication, inform our expectations about potential future outcomes in ways that less theoretically based models, such as gravity models, could not. Additionally, they can be adjusted extensively to better represent the real system they simulate. As a consequence, we have the opportunity to learn more through "dialogue with the model" than has been true in the past, and by understanding the limits of the model via uncertainty metrics, we better understand what conclusions we may draw. At PSRC, forecast confidence intervals have already made a contribution in this regard.

The ability to use the model system as a learning tool is no better embodied than in scenario simulations, the next major planned application for UrbanSim at PSRC. The 2013 Land Use Baseline Forecast gives information about the likelihood of future outcomes, given existing policies. Scenario simulations could characterize how the probability of a future land-use outcome might change as a result of altered policy, for example.

In this paper, we demonstrate how a methodology for assessing uncertainty in land-use models developed in an academic environment was adopted into practical use by a public planning agency and (to our best knowledge for the first time) used in an official land use forecast. The methodology itself is not specific to land-use models; it can be applied to any simulation models that fit into the framework illustrated in Figure 2. We hope this work contributes to the emerging efforts to incorporate uncertainty into the forecasting process by providing a feasible option for other practitioners to draw from in their own implementations.

References

Brown, D. G., S. Page, R. Riolo, M. Zellner, and W. Rand. 2005. Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science*, 19(2):153–174.

de Beer, J. and M. Alders. 1999. *Probabilistic population and household forecasts for the Netherlands*. Statistics Netherlads, Department of Population. URL http://www.cbs.nl/nr/rdonlyres/5b3209cc-92b4-4de8-b0fa-0b30ceaad427/0/probabilisticforecasts.pdf.

Duthie, J., A. Voruganti, K. M. Kockelman, and S. T. Waller. 2010. Uncertainty analysis and its impact on decision-making in an integrated transportation and gravity-based land use model. *Journal of Urban Planning and Development*, 136(4):294–302.

Felsenstein, D., K. Axhausen, and P. Waddell. 2010. Land use-transportation modeling with UrbanSim: Experiences and progress. *Journal of Transport and Land Use*, 3(2).

Gerland, P., A. E. Raftery, H. Ševčíková, N. Li, D. Gu, T. Spoorenberg, L. Alkema, B. K. Fosdick, J. L. Chunn, N. Lalic, G. Bay, T. Buettner, G. K. Heilig, and J. Wilmoth. 2014. World population stabilization unlikely this century. *Science*, 346:234–237.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In P. Zarembka, ed., *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

McFadden, D. 1978. Modeling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Wiebull, eds., *Spatial Interaction Theory and Planning Models*, pp. 75–96. Amsterdam: North Holland.

McFadden, D. 1981. Econometric models of probabilistic choice. In C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*, pp. 198–272. Cambridge, MA: MIT Press.

Pontius Jr, R. G. and J. Spencer. 2005. Uncertainty in extrapolations of predictive land-change models. *Environment and Planning B: Planning and Design*, 32:211–230.

Poole, D. and A. E. Raftery. 2000. Inference for deterministic simulation models: the Bayesian melding approach. *Journal of the American Statistical Association*, 95(452):1244–1255.

Pradhan, A. and K. M. Kockelman. 2002. Uncertainty propagation in an integrated land use-transportation modeling framework - output variation via UrbanSim. *Transportation Research Record: Journal of the Transportation Research Board*, 1805(1):140–150.

Puget Sound Regional Council. 2013a. *2013 Forecast Products*. URL http://www.psrc.org/data/forecasts/2013-forecast-products.

Puget Sound Regional Council. 2013b. *2013 Land Use Baseline Validation Report*. URL www.psrc.org/assets/9876/ForecastProdBaselineValidation.pdf.

Raftery, A. E., G. H. Givens, and J. E. Zeh. 1992. Inference from a deterministic population dynamics model about bowhead whale, *balaena mysticetus*, replacement yield. Technical Report 232, Department of Statistics, University of Washington, Seattle, Wash.

Raftery, A. E., G. H. Givens, and J. E. Zeh. 1995. Inference from a deterministic population dynamics model for bowhead whales. *Journal of the American Statistical Association*, 90(430):402–416.

Raftery, A. E., N. Li, H. Ševčíková, P. Gerland, and G. K. Heilig. 2012. Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences*, 109:13915–13921.

Rodier, C. J. 2005. *Verifying the Accuracy of Land Use Models Used in Transportation and Air Quality Planning: A Case Study in the Sacramento, California Region*. Institute of Transportation Studies, University of California Davis. URL transweb.sjsu.edu/MTIportal/research/publications/summary/0502.html.

Rodier, C. J. and R. A. Johnston. 2002. Uncertain socioeconomic projections used in travel demand and emissions models: could plausible errors result in air quality nonconformity? *Transportation Research Part A: Policy and Practice*, 36(7):613–631.

Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–null. doi: 10.1086/260169.

Ševčíková, H., A. Raftery, and P. Waddell. 2007. Assessing uncertainty in urban simulations using Bayesian melding. *Transportation Research B*, 41:652–669.

Ševčíková, H., A. Raftery, and P. Waddell. 2011. Uncertain benefits: Application of Bayesian melding to the Alaskan Way Viaduct in Seattle. *Transportation Research A*, 45:540–553.

Tayman, J. 1996. The accuracy of small-area population forecasts based on a spatial interaction land use modeling system. *Journal of the American Planning Association*, 62(1):85–98. doi: 10.1080/01944369608975672.

Tayman, J. 2011. Assessing uncertainty in small area forecasts: State of the practice and implementation strategy. *Population Research and Policy Review*, 30:781–800. doi: 10.1007/s11113-011-9210-9.

United Nations. 2013. *World Population Prospects: The 2012 Revision*. New York, NY: United Nations.

Waddell, P. 2002. UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association*, 68(3):297–314.

Waddell, P. 2010. Modeling residential location in UrbanSim. In J. Preston, F. Pagliara, and D. Simmonds, eds., *Modelling Residential Location Choice*, pp. 165–180. Springer.

Waddell, P. 2011. Integrated land use and transportation planning and modeling: Addressing challenges in research and practice. *Transport Reviews*, 31(2):209–229.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson. 2003. Microsimulation of urban development and location choices: Design and implementation of UrbanSim. *Networks and*

Spatial Economics, 3(1):43–67.

Waddell, P., H. Ševčíková, D. Socha, E. Miller, and K. Nagel. 2005. Opus: An Open Platform for Urban Simulation. Presented at the Computers in Urban Planning and Urban Management Conference, London. Available from www.urbansim.org/papers.

Waddell, P., G. Ulfarsson, J. Franklin, and J. Lobb. 2007. Incorporating land use in metropolitan transportation planning. *Transportation Research Part A: Policy and Practice*, 41:382–410.

Waddell, P., L. Wang, B. Charlton, and A. Olsen. 2010. Microsimulating parcel-level land use and activity-based travel: Development of a prototype application in San Francisco. *Journal of Transportation and Land Use*, 3(2):65–84.

Wang, L., P. Waddell, and M. Outwater. 2011. Incremental integration of land use and activity-based travel modeling: Work-at-home and out-of-home workplace choices. Transportation Research Board 89th Annual Meeting Compendium of Papers DVD.

Zhao, Y. and K. M. Kockelman. 2002. The propagation of uncertainty through travel demand models: An exploratory analysis. *Annals of Regional Science*, 36(2):145–163.