PROJECT DOCUMENT

County of Hawai'i General Plan Comprehensive Review

Trend Scenario and Land Use Allocation Technical Report

Date: April 18, 2016
Authors: Amy DeBay, Ian Varley, Doug Walker

Introduction

Task S2 (Land Use Allocation) sets up a framework for estimating future development patterns (amounts and location) based on a set of rules. Task E (Trend Scenario) uses an initial, calibrated run of the allocation model to estimate future development patterns based on historical trends. These tasks go hand-in-hand and are combined in this technical report.

Allocation Concepts and Approach

Future development patterns (amounts and location) are estimated using an algorithm-driven process called allocation. Allocation models the interplay between market demand for development in certain locations ("desirability") and amount of development allowed according to current regulations or by future land use patterns suggested by alternate scenarios ("capacity"). Given a pre-determined amount of growth expected in the given time frame (here, new growth between 2015 and 2040), the allocation process estimates where each incremental unit of new development will go, following the basic presumption that the most desirable areas will be developed first, capacity allowing. Thus highly desirable areas are assigned growth first, and then slightly less desirable areas get developed next, etc., until all the estimated growth amount has been accommodated. Numerous refinements to the basic principle are used to produce the final estimates. For example, parcels aren't always filled all the way to capacity, a certain amount of controlled "randomness" is often applied to the growth pattern, etc. For mixed use areas, both residential and non-residential growth can be assigned.

In this study, a CommunityViz tool called *Allocator 5* is used. The methods combined with the algorithms in Allocator 5 provide a well-reasoned analysis that will be helpful for this and myriad other planning studies, but it is recognized that the results have limitations in terms of modelling precision and confidence. The CommunityViz allocation method is sometimes described as "light-weight" to "medium-weight," differentiating it from the "heavy-weight" algorithms such as UrbanSim or PECAS that are considerably more sophisticated but are more difficult and expensive to implement. In contrast, the Allocator 5 method is easier to use and lends itself well to "what-if" type scenario planning.

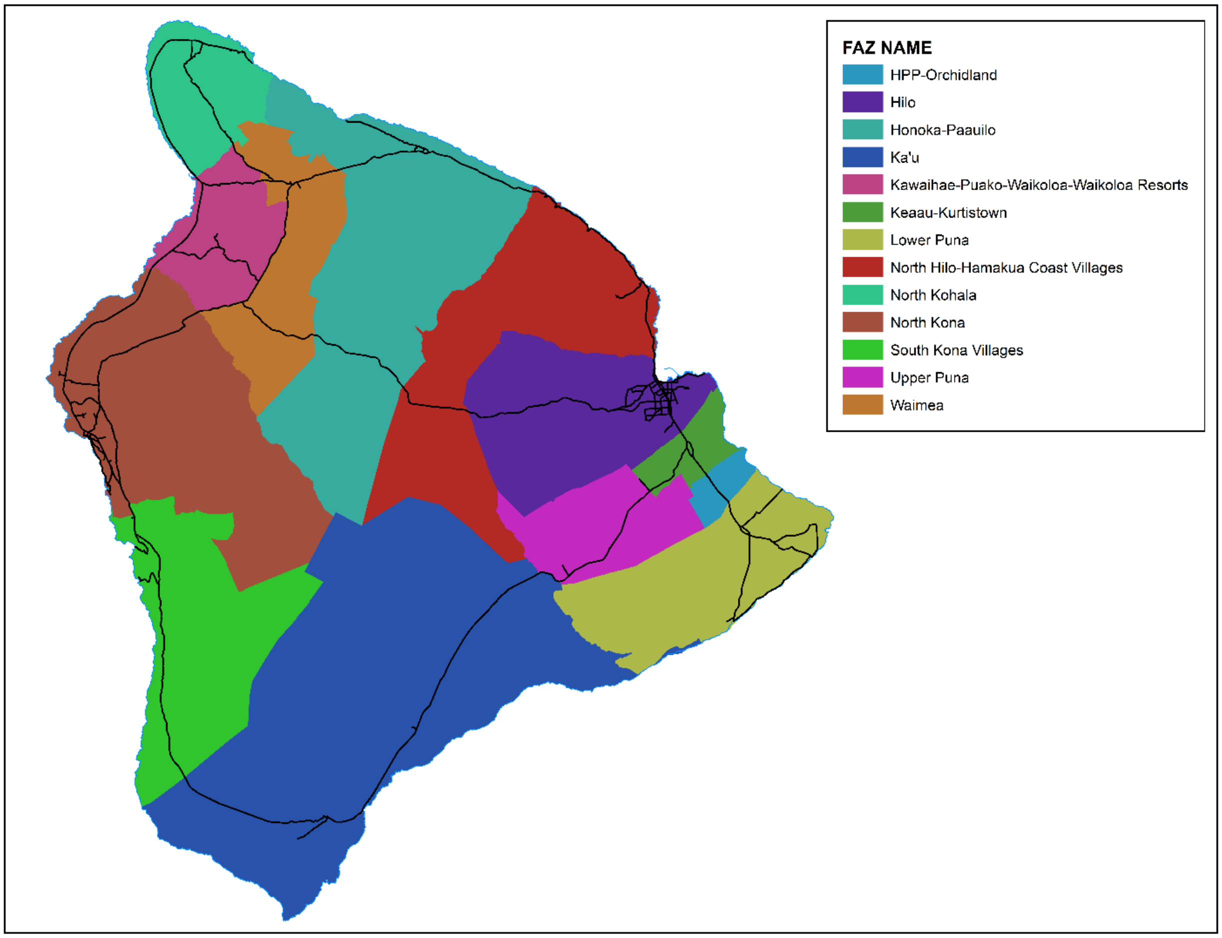
At the highest level, the allocation algorithm takes 3 inputs—growth amount, capacity, and desirability—and generates 1 output—a pattern of future development. Our methods for each of these are described next.

Growth Amount

As a 2015 baseline for housing unit data, Placeways used data from the County's Real Property Tax (RPT) office to identify the number of housing units and non-residential square feet. The procedure began with a database file from RPT that, unfortunately, lacked metadata, so the fields were interpreted manually. For each TMK, buildings were converted to housing units where appropriate (buildings and dwelling units were tracked separately). Single family homes and ʻohana units were readily identified in the RPT data. Multifamily housing required some additional steps. In cases where there were multiple records per TMK (as with condos), the records required consolidation to identify the total number of dwelling units per TMK. Separate analyses were performed for single family, multifamily, and commercial properties. In RPT data, housing units that are treated as commercial property (e.g., apartments and timeshares) were counted as residential only when the RPT data showed them as such. Once this basic processing was complete, Placeways used the RPT online tool, Google Street View, digital air photos and other tools to verify the number of existing units with the goal of establishing an accurate baseline and using the RPT database to its maximum extent.

For the trend scenario, growth projections are provided by SMS, a Hawaiʻi-based research and consulting company (see their report "General Plan Comprehensive Review Trends and Forecast Analysis Final Report (2015)"). These projections are broken out by 13 geographic areas called "forecast analysis zones" or FAZs (see Figure 1) and by use type (residential dwelling units and non-residential square feet). In order to add additional land use information to the allocation, the SMS forecasts were further broken down into four categories: single family dwelling units, multifamily dwelling units, commercial square feet, and industrial square feet. The ratio of single family to multifamily was found using the mean of the ratio from three dates in the recent past (2000, 2010, 2015) for which the ratio was known. This ratio was then applied to the combined residential allocation amounts to produce the single family/multifamily splits seen in Tables 1, 3, and 6. Similarly, growth amounts for non-residential development were developed as a single forecast and had to be split into amounts for commercial and industrial uses. The 2013 ACS Employment by Occupation Type data were used to identify the ratio of industrial employment to commercial employment and to produce the splits seen in Tables 2, 3, and 7. While this method assumes no change in the ratio of single family to multifamily units and commercial to industrial space, it reflects the recent development patterns in the Trend Scenario and can easily be adjusted for use in alternate scenarios.

Figure 1. Map of Forecast Analysis Zone (FAZ) Areas (source: SMS)



In contrast to the RPT-based method for establishing a baseline, the housing forecast data produced by SMS rely on US Census and Hawaiʻi DBEDT as sources. The methods used to collect Census and DBEDT data are quite different from RPT, resulting in differing 2015 baseline quantities. In addition, SMS did not suggest a 2015 baseline, instead using decadal increments for forecasting. In order to establish an SMS 2015 baseline, Placeways used the average of SMS's 2010 and the first forecast year of 2020. This results in a 2015 baseline difference of 11,558 housing units (RPT 2015: 75,100; SMS/Census: 86,658). The reasons that the RPT and Census derived baseline amounts are different are due to the sources' two distinct methods, and no attempt was made to reconcile them.

To calculate the amount of new residential growth, Placeways used a method to find the *relative* amount of net new amount of growth per FAZ. This method finds the percent change, per SMS, between 2015 and 2040 and applies that to the RPT 2015 baseline. This forecast results in fewer net new units (35,750) than the SMS forecast (40,160 new units), but its rates of change match SMS.

The SMS non-residential forecasts were already reconciled with the RPT data, and their forecast was calibrated to closely match the 2015 RPT baseline square footage. Therefore, in the case of non-residential growth, there was no need to rectify the forecast numbers as was the case with residential growth. Non-residential square feet were rounded and translated from square feet to 1,000 square feet for the purposes of allocation. This ensured that allocation amounts are in whole increments and not in very small portions of square feet.

Table 1. Residential Forecast by FAZ

FAZ	2015			2040 Change		2040 Total		
	Single Family	Multi-family	Total	Single Family	Multi-family	Single Family	Multi-family	Total
Hilo	14,713	1,138	15,851	2,953	833	17,666	1,971	19,636
North Hilo - Hāmākua Coast Villages	2,822	12	2,834	721	71	3,543	83	3,626
Honokaʻa-Paʻauilo	2,399	14	2,413	871	56	3,270	70	3,340
Waimea	3,212	98	3,310	1,420	158	4,632	256	4,887
North Kohala	2,499	17	2,516	785	59	3,284	76	3,360
Kawaihae-Puakō-Waikoloa-Waikoloa Resorts	2,610	3,390	6,000	1,337	1,337	3,947	4,727	8,675
North Kona	11,181	5,989	17,170	4,418	2,708	15,599	8,697	24,295
South Kona Villages	3,437	73	3,510	1,129	125	4,566	198	4,765
Kaʻū	3,397	76	3,473	2,135	112	5,532	188	5,720
Keaʻau -Kurtistown	1,640	10	1,650	834	35	2,474	45	2,518
Upper Puna	4,884	0	4,884	3,373	104	8,257	104	8,361
HPP-Orchidland	6,654	0	6,654	7,431	152	14,085	152	14,237
Lower Puna	4,835	0	4,835	2,515	78	7,350	78	7,428
Total	**64,283**	**10,817**	**75,100**	**29,922**	**5,828**	**94,205**	**16,645**	**110,850**

Table 2. Non-Residential Forecast by FAZ (in 1,000 square feet)

FAZ	2015			2040 Change		2040 Total		
	Commercial	Industrial	Total	Commercial	Industrial	Commercial	Industrial	Total
Hilo	9,187	3,762	12,949	3,979	346	13,166	4,108	17,274
North Hilo - Hāmākua Coast Villages	347	30	377	98	8	445	38	483
Honokaʻa-Paʻauilo	438	38	476	130	7	568	45	613
Waimea	1,303	75	1,379	421	37	1,724	112	1,837
North Kohala	290	255	545	181	12	471	267	738
Kawaihae-Puakō-Waikoloa-Waikoloa Resorts	5,406	63	5,470	1,608	85	7,014	148	7,163
North Kona	6,512	5,135	11,648	3,400	296	9,912	5,431	15,344
South Kona Villages	868	16	884	235	18	1,103	34	1,137
Kaʻū	303	0	303	110	11	413	11	424
Keaʻau-Kurtistown	760	902	1,662	454	51	1,214	953	2,167
Upper Puna	201	2	203	54	3	255	5	260
HPP-Orchidland	120	0	120	57	4	177	4	181
Lower Puna	413	0	413	114	14	527	14	541
Total	**26,150**	**10,279**	**36,428**	**10,841**	**892**	**36,991**	**11,171**	**48,161**

Capacity

Capacity values used in this report are based primarily on the results of Task S1, covered in a separate report. The focus is on net capacity, which is gross (or total) capacity minus existing development. Numeric capacity is assigned to every parcel on the island for residential dwelling units, and a separate numeric capacity for non-residential square feet. The majority of visitor units, existing and projected, are within 3 of the 13 FAZ areas. More information on visitor units and how they are addressed can be found in the Indicator Modeling technical report (Task G).

During the initial calibration of the allocation model, it was decided to place caps on the capacity of individual parcels that limited the amount of development that could occur on any single parcel. Caps for both residential and non-residential development were identified by looking at the County's subdivision records and the 2004-2015 building permit data. This information revealed what the recent historical maximums were for each kind of development: 600 units for residential development and 306,000 square feet for non-residential development. (The historical maximums and their place in the model are also discussed below in the section on additional calibration factors.) The final numbers below represent capped net capacity. Later in the project, these capacity values may be adjusted to model different potential planning and policy decisions.

Table 3. Net Capacity (Capped) by FAZ

FAZ Name	Residential Single Family Capacity (DU)	Residential Multifamily Capacity (DU)	Commercial Capacity (sq ft)	Industrial Capacity (sq ft)
Hilo	22,299	1,183	7,037,091	6,766,170
North Hilo - Hāmākua Coast Villages	8,721	50	1,173,273	894,155
Honokaʻa-Paʻauilo	9,691	39	579,692	133,073
Waimea	6,807	827	1,242,247	577,334
North Kohala	9,950	231	1,660,320	1,499,805
Kawaihae-Puakō-Waikoloa-Waikoloa Resorts	15,237	3,815	1,944,304	310,342
North Kona	21,855	5,376	9,831,956	12,956,188
South Kona Villages	21,445	190	539,714	0
Kaʻū	25,088	1,971	768,304	620,075
Keaʻau-Kurtistown	5,518	75	1,481,939	4,484,680
Upper Puna	25,390	31	135,975	0
HPP-Orchidland	9,996	0	0	0
Lower Puna	25,928	151	656,727	74,451
Total	**207,925**	**13,939**	**27,051,542**	**28,316,273**

Desirability

Desirability is a complex topic and represented a large proportion of the effort for this task.

The desirability of a given parcel for a given use (residential or commercial) is represented by a score from 0 (least desirable) to 100 (most desirable). A parcel that is not eligible for a given use is assigned a desirability score of -1. On a map, the pattern of desirability scores is sometimes called a "desirability surface" because one can picture a lumpy blanket covering the island with high points in areas of high desirability and low points in areas of low desirability.

Desirability was calculated all at once for the entire island. Given more time and resources, it would be possible to perform separate calculations for subareas such as east/west or individual FAZs. However, the additional effort would make little difference because growth amounts are constrained to meet FAZ-specific totals and thus growth is allocated based on relative desirability scores within a FAZ, not between FAZs.

The detailed procedure for creating the initial desirability surface, used to create the Trend Scenario, follows:

1. Spatial and non-spatial factors affecting location desirability were hypothesized. These factors, such as proximity to infrastructure or coastline, likely affected development desirability in the past and could be quantified for all parcels on the island given available information. While a complete list of factors tested is included at the end of this report, all hypothesized factors could fit into the following categories:
    - Distance to infrastructure, geographical features, town and commercial centers
    - Parcel shape/size
    - Terrain/climate characteristics
    - Neighbor (10 ft) and neighborhood (1000 ft) context: the number of nearby parcels and development proximity (the number of nearby parcels that are built)
    - Current land use and build status
    - Geographic location (District, FAZ, etc.)
    - Property and building valuation.

    Using CommunityViz, these factors were calculated for each parcel on the island and exported to a table for use in SPSS statistical analysis software.

2. Statistical regression analysis in SPSS was used to calculate how well each factor correlates with new development in a given time frame. Three timeframes were initially considered:
    - All development since the beginning of RPT (County of Hawaiʻi Real Property Tax Office) recordkeeping (1880s)
    - All development since 1995
    - All development between 1975 and 1995.

    However, because the goal of this statistical regression was to capture the principal factors influencing urban growth in the *recent* past, a cutoff date of 1995 was established. This year was selected to encompass a full cycle of real estate development and to capture the "highs and lows" in development activity. Development during this period followed this approximate pattern:

    - Trough: 1996
    - Peak: 2005-2006
    - Trough: 2009-2011
    - Recent uptick: 2015.

    The SMS CoH 2016 General Plan Final Report (Figure 1, Resident Population) illustrates this pattern.

    The statistical analysis models the relationship between dependent variables – Commercial Square Feet (COM_SQFT) and Number of Residential Buildings (NumberOfResidences) – and the hypothesized list of independent variables (see Appendix 1, Table 12. Factors Tested for Historical Growth Trends). As a result, a step-wise Multiple Linear Regression model was chosen to create coefficients associated with each of the variables in order to represent the independent contributions of each independent variable to the prediction of the dependent variable after controlling for all other independent variables.

The initial analysis included an overall view of development, an earlier era of development (1975-95), and post-1995 recent development patterns for both of the dependent variables (see Appendix 1: *Comparing Post 1995 Regression Factors with the 1975-1995 Regression Factors* for a discussion of these results). The results show the top 10 variables—that is, the top 10 of the hypothesized desirability factors—that influence each of the analyses, along with the absolute value of each of the standardized coefficients. The coefficient values allow ranking the variables from most to least influential. Detailed analysis information is included at the end of this report.

The standardized regression coefficients with the 10 highest absolute beta values for the post-1995 period were converted into CommunityViz weighting factors normalized to the scale 0 – 10, where 0 is no correlation and 10 is the highest correlation of any factor (though less than 1). Candidate factors with lower beta coefficients, below the top ten, were ignored for the desirability score.

A cutoff of 10 factors was chosen for a few reasons. One was to keep the most significant factors in the mix. The top 10 account for the majority of the causal influence of all factors tested. Additionally, there was a benefit to limiting the number of movable parts for testing. Even with 10, it is challenging to understand the interplay of all inputs and the individual effect on the overall score. The goal here is to incorporate both sophistication and manageable interactive parts, and 10 is a reasonable number for that.

Some factors are negatively correlated, and some factors are inversely correlated. For instance, distance values that correlate to growth are often inverse: nearer, smaller distance values are more desirable and further, larger distance values are less desirable. In the table below, negative and inverse correlations are indicated by a negative beta value score. Understanding this, many factors below make intuitive sense. The slope factor for example suggests that as land gets steeper (slope increases), the likelihood of development decreases. Some results are not always intuitive, however. Statistically, for example, it is found that parcels that are *closer* to old lava flows are more desirable for non-residential development than those far away (i.e. Distance2LavaFlow1790). Some positively correlated cases also benefit from some explanation. For example, the strongest factor for residential development is Distance2VolcanoHazard. This is a positively correlated factor meaning that as distance increases away from volcano hazards, the likelihood of development also increases.

Table 4. Residential Factors (Top 10) for Growth Desirability Model (Post 1995 Development)

Factor	Description	Beta Value	Absolute Beta Value	Normalized Value (Weighting)
Distance2VolcanoHazard	Distance to high volcano risk area, defined as areas classified as category 1 or 2 on the risk layer	0.0808	0.0808	10
Distance2LavaFlow1250	Distance to older lava flow, deposited between the year 1250 and the present	-0.0575	0.0575	7.1
Slope	Mean slope of the parcel in percent rise	-0.0494	0.0494	6.1
Intersections1miDensity	Road intersection density: the number of intersections on major roads within ½ mi of the parcel divided by area of the parcel	-0.0439	0.0439	5.4
Rainfall	Annual average rainfall, in inches	-0.0429	0.0429	5.3
Distance2CenterAdjusted	Network distance to commercial center. Commercial center is defined by RPT land use "commercial." This adjusted value uses straight line distance for those features not picked up by the patchy network.	-0.0411	0.0411	5.1
RESpost1995ProximitySum1000ft	Total number of dwelling units on parcels within 1000 ft of a parcel built after 1995	-0.0394	0.0394	4.9
RESpost1995ProximitySum10ft	Total number of dwelling units on parcels within 10 ft of a parcel built after 1995	-0.0321	0.0321	4
ParcelPerimeter2DivArea	Lot shape: perimeter squared divided by area	0.0305	0.0305	3.8
Distance2Coastline	Distance to the island's coastline	-0.0301	-0.0301	3.73

Table 5. Non-Residential Factors (Top 10) for Growth Desirability Model (Post 1995 Development)

Factor	Description	Beta Value	Absolute Beta Value	Normalized Value
Distance2LavaFlow1790	Distance to recent lava flow, deposited since 1790	-0.3410	0.3410	10
Intersections1miDensity	Road intersection density: the number of intersections on major roads within ½ mi of the parcel divided by area of the parcel	-0.2560	0.2560	7.5
Distance2Airports	Distance to nearest major airport	0.2246	0.2246	6.6
Proximity1000ftParcelDensity	The number of parcels within 1000ft divided by the area of the parcel	0.1723	0.1723	5
Distance2ExistingMF1	Distance to nearest multifamily residential development	-0.1380	0.1380	4
Distance2SewerService	Distance to nearest wastewater service line	-0.1207	0.1207	3.6
Distance2WaterService	Distance to nearest potable water service line	0.1089	0.1089	3.2
Proximity10ftParcelDensity	The number of parcels within 10 ft divided by the area of the parcel	-0.1006	0.1006	3
ParcelPerimeter2DivArea	Lot shape: perimeter squared divided by area	-0.0935	0.0935	2.7
Distance2VolcanoHazard	Distance to high volcano risk area, defined as areas classified as category 1 or 2 on the risk layer	0.0885	0.0885	2.6

3. The CommunityViz Suitability Wizard was then used to create a suitability analysis using the parcels' normalized factor values (such as distance to roads) as inputs. Since these values were pre-calculated for the statistical analysis, Suitability Wizard was pointed directly at the numeric values, improving processing performance. The wizard default is set to 5 (on a scale of 0 to 10) for each factor weighting. After the suitability analysis is run with defaults, the weighting assumption defaults are set for each factor according to the values calculated in the previous step to calibrate the score to historical trends.
4. The suitability (desirability) score for each parcel is displayed using the parcels layer symbolized by the suitability scores. The factor weightings are adjustable, so each factor that contributes to the overall score can be given a level of priority appropriate to the goals of the scenario.

Trend Scenario Allocation

The allocation of forecasted development in the "Trend Scenario" is based on the desirability surface representing historic development trends. Later in the project, it will be possible to develop alternate desirability surfaces for scenario planning in which planners test planning strategies that could encourage growth to evolve in particular ways.

Allocation is performed using Allocator 5 with the inputs above. Again, allocation distributes predetermined growth amounts for each FAZ to parcels within that FAZ, developing the most desirable parcels first and proceeding until all growth has been accommodated. No parcel receives more development than it has capacity for, and many parcels receive little or no development even though they have capacity available. For the initial run, used to develop the Trend Scenario, randomness is set at 1 out of 10. Results for each FAZ are as follows:

Table 6. Allocated Residential Growth by FAZ

FAZ Name	Residential Single Family Dwelling Units	Remaining Single Family Capacity	Residential Multifamily Dwelling Units	Remaining Multifamily Capacity
Hilo	2,952	19,347	833	350
North Hilo - Hāmākua Coast Villages	721	8,000	50	0
Honokaʻa-Paʻauilo	871	8,820	39	0
Waimea	1,419	5,388	158	669
North Kohala	785	9,165	59	172
Kawaihae-Puakō-Waikoloa-Waikoloa Resorts	1,338	13,899	1,338	2,477
North Kona	4,418	17,437	2,708	2,668
South Kona Villages	1,130	20,315	126	64
Kaʻū	2,135	22,953	112	1,859
Keaʻau-Kurtistown	833	4,685	35	40
Upper Puna	3,373	22,017	31	0
HPP-Orchidland	7,431	2,565	0	0
Lower Puna	2,515	23,413	78	73
Total	**29,921**	**178,004**	**5,567**	**8,372**

Table 7. Allocated Non-Residential Growth by FAZ in 1,000 square feet

FAZ Name	Commercial Allocated	Remaining Commercial Capacity	Industrial Allocated	Remaining Industrial Capacity
Hilo	3,979	3,039	346	6,424
North Hilo - Hāmākua Coast Villages	98	1,074	8	887
Honoka'a-Pa'auilo	130	450	7	125
Waimea	421	814	37	541
North Kohala	181	1,479	12	1,487
Kawaihae-Puakō-Waikoloa-Waikoloa Resorts	1,608	336	85	224
North Kona	3,400	6,411	296	12,656
South Kona Villages	235	303	0	0
Ka'ū	110	655	11	609
Kea'au-Kurtistown	454	1,023	51	4,434
Upper Puna	54	80	0	0
HPP-Orchidland	0	0	0	0
Lower Puna	114	542	14	60
Total	**10,784**	**16,206**	**867**	**27,447**

Trend Calibration Factors

In order to better replicate the patterns of the previous 20 years, additional calibration factors were developed to supplement the factors used in the statistical analysis. The objective of the Trend Scenario is to reflect a continuation of recent patterns, and these factors help reflect the recent development characteristics. While the statistical analysis provides an impartial and "data-driven" perspective on the patterns of recent development, its results cannot provide a complete picture because of a number of limitations:

- Data limitations. Data were not available on all factors and across all historic time scales. In particular, the model lacked in-depth real estate market data, including market conditions and consumer preferences for both residential and non-residential development. In addition, historic building data were not available in appropriate forms.
- Geographic biases that resulted from using parcel data as the unit of analysis. For example, large parcels tended to score very well in the statistical analysis. These features tend to show advantages that small parcels do not: they have more capacity, they have more neighboring parcels, and they can be closer to more desirable locations simply because of their size. A grid-based analysis would reduce these issues, but it would introduce many other issues (e.g., mismatch between the grids and the available parcel data) in their place.
- Difficulty reflecting all patterns and preferences for all the island at a local scale (e.g., what drives growth in North Kona is somewhat different from what drives growth in Lower Puna). It was outside the scope of this project to run individual analyses for smaller areas (e.g., FAZs).

- No forecast model is perfect, and it is expected that any algorithmic results will require some degree of adjustment to match observed results.

Because of these limitations on factors derived solely from statistical analysis, additional calibration factors based on expert human knowledge were included in the analysis as well. The combination resulted in a hybrid system that carefully combines both statistical modeling and expert judgement. Each calibration factor works in a different way to help fine-tune the model by what is believed as common knowledge by County of Hawaiʻi planners and by comparing values observed in the recent development data to the results of the draft versions of the Trend Scenario allocation.

Some of these factors (Redevelopment Friction, Residential Subdivision Friction and Non-residential Size Friction) help steer growth towards areas that have received consistent development pressure in the previous 20 years. The Pipeline Projects Factor prioritizes parcels that currently have development proposals pending. While these factors help calibrate the Trend Scenario to reflect recent patterns, it is difficult—and not necessarily desirable—to exactly replicate rates and patterns of the recent past with what occurs in the Trend Scenario's modeled future. The future is unknown and many external factors and unforeseeable conditions will affect how growth actually occurs. The goal of the Trend Scenario is to provide a plausible and useful reference for comparing alternative scenarios to help inform policy and the contents of the General Plan; it is not intended as a detailed forecast.

Recent Lava Factor

While a significant amount of development in the County of Hawaiʻi occurs on geologically recent lava flow (often on lava less than a few hundred years old), lava that has been deposited very recently is a significant obstacle to development. Lava flows that are more recent than 1990 were identified, and desirability of parcels that overlapped post-1990 lava flows was reduced.

Pipeline Projects Factor

This factor adds a bonus to the desirability of parcels that are currently under development or that County planning staff assume are likely to develop in the near future. The pipeline projects are divided into two groups. Group one consists of two projects, Kamakana Villages and Kealakehe Homesteads, which are very likely to develop or are in the process of development. The second group includes five projects that are less advanced in the planning process but have greater than average chances of developing. The project pipeline projects are mostly located in North Kona and primarily entail residential development. Table 8 identifies the pipeline projects and the number of units allocated to them.

Table 8. Pipeline projects

Project Name	Approximate Location	Allocated Residential Units
Kamakana Villages	**Ane Keohokālole Highway, North of Kailua Kona**	**600**
Kealakehe Homesteads	**North of Kailua-Kona, Corner of Keanalehu & Manawalea**	**184**
UH Pālamanui	UH West Hawaiʻi campus, just east of airport	300
Pualani	South of Kailua-Kona along Queen Kaʻahumanu highway	178
Keahuolū Queen Liliʻuokalani Trust	Various parcels, north of Kailua-Kona downtown	48

Redevelopment Friction

The uncalibrated allocation model orients development towards areas of greater desirability and available capacity. Many parcels already have existing residential or non-residential structures on them but according to the capacity analysis have some additional capacity (these parcels are considered "underbuilt" for purposes of the model). The Redevelopment Friction factor considers redevelopment in the sense of any new development on a parcel that has an existing residential or non-residential structure (according to RPT). The uncalibrated model does not distinguish the nature of this new development: whether it is a physical addition to a structure, a new structure on an undeveloped portion of the parcel, or the wholesale redevelopment of an existing structure. Redevelopment areas can sometimes be more difficult and costly to develop but they may be highly desirable because of their location or other amenities. The 2004-2015 building permit data reveal that 34% of residential development occurred on parcels that already had a residential or non-residential structure. This figure is lower for commercial development, where 15% of growth occurred on parcels with existing structures. The Redevelopment Friction factor applies a penalty to the desirability score for both residential and non-residential development to parcels that already have some development.

Residential Subdivision Friction

An analysis of residential growth since 1995 revealed that over 75% of new dwelling units were built on parcels that were subdivided to sizes of less than 5 acres. Subdivisions yielding parcels larger than 25 acres were not as common and did not contain large amounts of residential development in the recent past. The County's subdivision data revealed that in the last fifteen years, the largest subdivision to be approved was 590 parcels. During initial runs of the allocation model, it was observed that the tool tended to select larger parcels over smaller parcels, a pattern that was not consistent with the recent development patterns. The County's subdivisions layer depicts all of the island's major subdivisions and using this data as a guide, subdivided parcels were flagged and received higher desirability scores than non-subdivided parcels. This factor actually consists of two factors, one factor that penalizes large parcels by size and another that caps residential development at 600 units, similar to the historic maximum.

Non-residential Footprint Friction

Similar to residential subdivision friction, recent non-residential development had certain characteristics related to building footprint size (i.e., the square footage of the non-residential development) that were not well captured by the regression analysis. Early runs of the model revealed some extremely large non-residential developments that seemed uncharacteristic with the recent development patterns. According to RPT records, the largest commercial development since 1995 was 306,000 square feet and the median size 6,510 square feet. For the Trend Scenario, non-residential capacity was capped at 300,000 and desirability was boosted for smaller capacity parcels.

Comparing Recent Development Trends and the Trend Scenario

After calibration, the Trend Scenario matches recent development patterns fairly well, as shown in the table below.

Table 9. Comparing recent development with the trend scenario

	Recent Development* 2004-2015	Trend Scenario
Residential Redevelopment	11% of recent growth occurred on built parcels	4% of growth occurs on built parcels
Commercial Redevelopment	15% of recent growth occurred on built parcels	16% of growth occurs on built parcels
Residential Subdivision Development	75% of recent growth occurred in existing subdivisions; median parcel size is 0.7 acres	73% of recent growth occurs in existing subdivisions; median parcel size is 0.9 acres
Non-residential Footprint Friction	Median development size was 6,510 square feet	Median development size is 5,000 square feet

* Recent redevelopment is based on RPT and building permit data; subdivision and parcel size factors use RPT data.

Appendix 1. Comparing Post 1995 Regression Factors with the 1975-1995 Regression Factors

The regression analysis conducted for all development and development post-1995 was also run for development that occurred between 1975 and 1995. Development in this era appeared to manifest some similar patterns (Distance to Volcano Hazards, Lava Flow, and Intersection Density) as the post 1995 development. Distance to centralized services such as water and sewer systems was more strongly correlated in the 1975-95 development, perhaps reflecting the closer proximity of development in that era to those services. In a more extreme case, the correlation for Distance2Airport reverses: it is negative for 1975-95 development and positive for post 1995 development. In other words, as distance to airports diminishes the likelihood of development increases in the 1975-95 era. The opposite is true for development in the post 1995 era.

Valuation factors (building, land values) were poorly correlated in the 1975-95 era for residential development. Valuation factors were more strongly correlated with non-residential square footage. However, the valuation factors used reflect present-day conditions. In general, care should be taken when comparing the results of the 1975-95 era beta values with the post 1995 beta values because this analysis relies on data that represents conditions as they are now, not as conditions were during the 1975-95 phase of the island's development. Some factors rely on data that has remained relatively consistent through time; for example the location of the island's coastline, slope, rainfall patterns, etc. are very similar now to how they were 40 years ago. Other factors have changed considerably since the 1975-95 development occurred (distance to commercial development, development proximity, land and building valuations, etc.). The 1975-95 development patterns would be better compared to the post 1995 development by factoring in the conditions as they existed in that 1975-95 era, not conditions as they exist in the present day. Unfortunately that level of analysis was not feasible within the scope of the current project. Below are two tables comparing the 1975-95 and the post 1995 beta values for residential and nonresidential development.

Table 10. Comparing post 1995 and 1975-95 non-residential beta values

Factor	ComPost95Beta	Com75-95Beta	Difference
Distance2LavaFlow1790	-.341	.006	-0.3474
Intersections1miDensity	-.256	-.100	-0.1564
Distance2Airports	.225	-.145	0.3700
Proximity1000ftParcelDensity	.172	-.042	0.2140
Distance2ExistingMF1	-.138	-.126	-0.0120
Distance2SewerService	-.121	.161	-0.2813
Distance2WaterService	.109	.013	0.0958
Proximity10ftParcelDensity	-.101	.019	-0.1196
ParcelPerimeter2DivArea	-.093	.077	-0.1703
Distance2VolcanoHazard	.088	.090	-0.0016

Table 11. Comparing post 1995 and 1975-95 residential beta values

Factor	ResPost95Beta	ResPost75-95Beta	Difference
Distance2VolcanoHazard	.081	.079	0.0017
Distance2LavaFlow1250	-.057	-.173	0.1151
Slope	-.049	-.049	0.0000
ValueBuilding	.049	.010	0.0389
Intersections1miDensity	-.044	-.078	0.0338
Rainfall	-.043	-.007	-0.0359
Distance24CenterAdjusted	-.041	-.028	-0.0132
RESpost1995ProximitySum1000ft*	-.039	#N/A	-0.0419
RESpost1995ProximitySum10ft**	-.032	#N/A	-0.0273
ParcelPerimeter2DivArea	.031	.037	-0.0067

* The 1975-95 results used the factor, RESpost7595ProximitySum1000ft instead of the post 1995 proximity sums, yielding the beta value of 0.002.

** The 1975-95 results used the factor, RESpost7595ProximitySum10ft instead of the post 1995 proximity sums, yielding the beta value of -0.005.

Table 12. Factors Tested for Historical Growth Trends

Factor Name	Description	Unit	Source
Acres	Size of the parcel in acres	acres	
Distance2Airports	Distance to nearest major airport	feet	County GIS
Distance2Belt10	Distance to the island's belt road	feet	County GIS
Distance2Coastline	Distance to coastline	feet	County GIS
Distance24CenterAdj	Network distance to commercial center. Commercial center is defined by RPT land use "commercial," this adjusted value used straight line distance for those features not picked up by the patchy network.	feet	RPT and county GIS
Distance2ElectricService	Distance to nearest electric utility pole	feet	County GIS
Distance2ExistingMF1	Distance to nearest multifamily residential development	feet	RPT and county GIS
Distance2ExistingSF1	Distance to nearest single family residential development	feet	RPT and county GIS
Distance2Hospitals	Distance to hospitals	feet	County GIS
Distance2LavaFlow1250	Distance to older lava flow, deposited since 1250	feet	USGS Hawaiʻi Geologic Map
Distance2LavaFlow1790	Distance to recent lava flow, deposited since 1790	feet	USGS Hawaiʻi Geologic Map
Distance2Major10	Distance to nearest major, arterial-type road	feet	County GIS

Factor Name	Description	Unit	Source
Distance2Schools	Distance to nearest public school	feet	County GIS
Distance2SewerService	Distance to nearest wastewater service line	feet	County GIS
Distance2Towns	Distance to nearest major town, towns defined by county provided "towns" layer	feet	County GIS
Distance2VolcanoHazard	Distance to high volcano risk area, defined as areas classified as category 1 or 2 on the risk layer	feet	USGS, via county GIS
Distance2WaterService	Distance to nearest potable water service line	feet	County GIS
Slope	Mean slope of the parcel in percent rise	percent rise	USGS DEM
ValueLand	Land value in dollars per the County's RPT records	dollars	RPT
Rainfall	Annual average rainfall, in inches	inches	University of Hawai'i
District	Planning district		County GIS
FAZ	Forecast analysis zone		SMS
MaukaMakaiNum	Uses 1 or makai (outside the belt) and 2 for mauka (inside the belt)		Placeways
COM_YRBLT_YN	If built 1, if not 0.		RPT
RES_YRBLT_YN	If built 1, if not 0.		RPT
isCom	If commercial, 1; if not, 0		RPT
isSF	If single family residential, 1; if not, 0		RPT
isMF	If multifamily residential, 1; if not, 0		RPT
COMAllYearsProximitySum10ft	Total commercial square feet on parcels within 10ft of any parcel with commercial square footage		RPT
COMAllYearsProximitySum1000ft	Total commercial square feet on parcels within 1000ft of any parcel with commercial square footage		RPT
COMpost1995PS1000ft	Total commercial square feet on parcels within 1000ft of any parcel with a structure built after 1995		RPT
COMpost1995PS10ft	Total commercial square feet on parcels within 10ft of any parcel with a structure built after 1995		RPT
RESAllYearsProximitySum10ft	Total number of dwelling units on parcels within 10ft of any parcel with a residential structure		RPT
RESAllYearsProximitySum1000ft	Total number of dwelling units on parcels within 1000ft of any parcel with residential structure		RPT
DUpost1995PS10ft	Total number of dwelling units on parcels within 10 ft of any parcel with a structure built after 1995		RPT
DUpost1995PS1000ft	Total number of dwelling units on		RPT

Factor Name	Description	Unit	Source
	parcels within 1000 ft of any parcel with a structure built after 1995		
ParcelPerimeter2DivArea	Lot shape: perimeter squared divided by area	square feet/feet	County GIS
Intersections1/2miDensity	Road intersection density: the number of intersections on major roads within ½ mi of the parcel divided by area of the parcel	intersection per acre	County GIS
Proximity10ftParcelDensity	The number of parcels within 10ft divided by the area of the parcel	Parcels per acre	County GIS
Proximity1000ftParcelDensity	The number of parcels within 1000ft divided by the area of the parcel	Parcels per acre	County GIS

Appendix 2. Detailed Statistical Methods

Commercial All Data

The first multiple linear regression was calculated to predict *COM_SQFT* for all possible parcels based on *ParcelPerimeter2DivArea*, *Intersections1miDensity*, *Slope*, *Rainfall*, *Distance2Major10*, *Distance2WaterService*, *Distance2SewerService*, *Distance2Schools*, *Distance2Coastline*, *Distance2VolcanoHazard*, *Distance2Airports*, *Distance2LavaFlow1250*, *Distance2LavaFlow1790*, *Distance2Towns*, *Distance2ElectricService*, *Distance2Hospitals*, *Distance24CenterAdjusted*, *Distance2ExistingSF1*, *Distance2ExistingMF1*, *Distance2Belt10*, *COMAllYearsProximitySum10ft*, *COMAllYearsProximitySum1000ft*, *ValueLand*, *Proximity10ftParcelDensity*, *Proximity1000ftParcelDensity* with n= 132,842. A significant regression equation was found ($F(28,132814) = 161.888$, $p < 0.001$) with an R^2 of 0.033.

After the analysis the 12 coefficients that influence Commercial square feet are listed below.

Table 13. Twelve most influential variables for commercial square feet, all years

Variables	Beta Value Coefficient
COMAllYearsProximitySum10ft	.125
COMAllYearsProximitySum1000ft	.056
Distance2VolcanoHazard	.054
Intersections1miDensity	-.051
Distance2Hospitals	.037
Distance2LavaFlow1790	.032
Distance2ExistingSF1	.025
Distance2ElectricService	-.023
Distance2ExistingMF1	-.020
Slope	-.019
Distance2SewerService	-.018
Distance2LavaFlow1250	-.018

Commercial Post 1995

The first multiple linear regression was calculated to predict *COM_SQFT* commercial buildings built on or after 1995 based on *ParcelPerimeter2DivArea*, *Intersections1miDensity*, *Slope*, *Rainfall*, *Distance2Major10*, *Distance2WaterService*, *Distance2SewerService*, *Distance2Schools*, *Distance2Coastline*, *Distance2VolcanoHazard*, *Distance2Airports*, *Distance2LavaFlow1250*, *Distance2LavaFlow1790*, *Distance2Towns*, *Distance2ElectricService*, *Distance2Hospitals*, *Distance24CenterAdjusted*, *Distance2ExistingSF1*, *Distance2ExistingMF1*, *Distance2Belt10*, *COMpost1995ProximitySum1000ft*, *COMpost1995ProximitySum10ft*, *ValueLand*, *Proximity10ftParcelDensity*, *Proximity1000ftParcelDensity* with n= 451. A significant regression equation was found ($F(27,424) = 2.259$, $p < 0.001$) with an R^2 of 0.126.

After the analysis the 12 coefficients that influence Commercial square feet after 1995 are listed below.

Table 14. Twelve most influential variables for commercial square feet, post 1995

Variables	Absolute Value Coefficient
Distance2LavaFlow1790	-.341
Intersections1miDensity	-.256
Distance2Airports	.225
Proximity1000ftParcelDensity	.172
Distance2ExistingMF1	-.138
Distance2SewerService	-.121
Distance2WaterService	.109
Proximity10ftParcelDensity	-.101
ParcelPerimeter2DivArea	-.093
Distance2VolcanoHazard	.088
ValueLand	.082
Distance2Towns	-.067

Commercial Post 1975-95

The first multiple linear regression was calculated to predict *COM_SQFT* commercial buildings built on or after 1975 and before 1995 based on *ParcelPerimeter2DivArea, Intersections1miDensity, Slope, Rainfall, Distance2Major10, Distance2WaterService, Distance2SewerService, Distance2Schools, Distance2Coastline, Distance2VolcanoHazard, Distance2Airports, Distance2LavaFlow1250, Distance2LavaFlow1790, Distance2Towns, Distance2ElectricService, Distance2Hospitals, Distance24CenterAdjusted, Distance2ExistingSF1, Distance2ExistingMF1, Distance2Belt10, COMpost1995ProximitySum1000ft, COMpost1995ProximitySum10ft, ValueLand, Proximity10ftParcelDensity, Proximity1000ftParcelDensity* with n= 687. A significant regression equation was found ($F(27,660) = 3.634$, $p < 0.001$) with an R^2 of 0.129.

After the analysis the 12 coefficients that influence Commercial square feet between 1975-95 are listed below.

Table 15. Twelve most influential variables for Commercial Square Feet > 1975 and < 1995

Variables	Beta Value Coefficient
Distance2SewerService	.161
Distance2Airports	-.145
ValueLand	-.144
Distance2ExistingMF1	-.126
Distance2ExistingSF1	-.126
Distance2Schools	.109
Intersections1miDensity	-.100
Distance2VolcanoHazard	.090
ParcelPerimeter2DivArea	.077
Distance24CenterAdjusted	-.072
Distance2Coastline	-.070
Distance2Belt10	-.059

Residential All Data

The first multiple linear regression was calculated to predict NumberOfResidences for residential buildings built after 1995 based on *ParcelPerimeter2DivArea, Intersections1miDensity, Slope, Rainfall, Distance2Major10, Distance2WaterService, Distance2SewerService, Distance2Schools, Distance2Coastline, Distance2VolcanoHazard, Distance2Airports, Distance2LavaFlow1250, Distance2LavaFlow1790, Distance2Towns, Distance2ElectricService, Distance2Hospitals, Distance24CenterAdjusted, Distance2ExistingSF1, Distance2ExistingMF1, Distance2Belt10, RESAllYearsProximitySum10ft, RESAllYearsProximitySum1000ft, ValueLand, Proximity10ftParcelDensity, Proximity1000ftParcelDensity* with n=132,842. A significant regression equation was found ($F(27,132,815) = 73.875$, $p < 0.001$) with an R^2 of 0.015.

After the analysis the 12 coefficients that influence Number of Residences the most are listed in the following table.

Table 16. Twelve most influential variables for residential, all years

Variables	**Beta Value Coefficient**
Distance2LavaFlow1250	-.062
Distance2VolcanoHazard	.060
Distance2ExistingMF1	-.050
Distance2SewerService	-.048
Intersections1miDensity	-.043
Distance2WaterService	.035
Distance2Coastline	-.028
Distance2Airports	.028
Rainfall	-.025
Distance2Towns	-.024
RESAllYearsProximitySum1000ft	.022
Slope	-.021

Residential Post 1995

The first multiple linear regression was calculated to predict NumberOfResidences for residential buildings built on or after 1995 based on *ParcelPerimeter2DivArea, Intersections1miDensity, Slope, Rainfall, Distance2Major10, Distance2WaterService, Distance2SewerService, Distance2Schools, Distance2Coastline, Distance2VolcanoHazard, Distance2Airports, Distance2LavaFlow1250, Distance2LavaFlow1790, Distance2Towns, Distance2ElectricService, Distance2Hospitals, Distance24CenterAdjusted, Distance2ExistingSF1, Distance2ExistingMF1, Distance2Belt10, DUpost1995ProximitySum10ft, DUpost1995ProximitySum1000ft, ValueLand, Proximity10ftParcelDensity, Proximity1000ftParcelDensity* with n=18,819. A significant regression equation was found ($F(27,18,792) = 11.549$, $p < 0.001$) with an R^2 of 0.016.

After the analysis the 12 coefficients that influence Number of Residences built after 1995 the most are found in the table below.

Table 17. Twelve most influential variables for residential, post 1995

Variables	Beta Value Coefficient
Distance2VolcanoHazard	.081
Distance2LavaFlow1250	-.057
Slope	-.049
Intersections1miDensity	-.044
Rainfall	-.043
Distance24CenterAdjusted	-.041
RESpost1995ProximitySum1000ft	-.039
RESpost1995ProximitySum10ft	-.032
ParcelPerimeter2DivArea	.031
Distance2Coastline	-.030
Distance2Schools	-.029
Distance2Hospitals	0.23

Residential 1975-95

The first multiple linear regression was calculated to predict NumberOfResidences for residential buildings built on or after 1975 and before 1995 based on *ParcelPerimeter2DivArea, Intersections1miDensity, Slope, Rainfall, Distance2Major10, Distance2WaterService, Distance2SewerService, Distance2Schools, Distance2Coastline, Distance2VolcanoHazard, Distance2Airports, Distance2LavaFlow1250, Distance2LavaFlow1790, Distance2Towns, Distance2ElectricService, Distance2Hospitals, Distance24CenterAdjusted, Distance2ExistingSF1, Distance2ExistingMF1, Distance2Belt10, DUpost1995ProximitySum10ft, DUpost1995ProximitySum1000ft, ValueLand, Proximity10ftParcelDensity, Proximity1000ftParcelDensity* with n=23,822. A significant regression equation was found ($F(27,23,795) = 17.956$, $p < 0.001$) with an R^2 of 0.020.

After the analysis the 12 coefficients that influence Number of Residences built between 1975-95 the most are found in the table below.

Table 18. Twelve most influential variables for Residential > 1975 and < 1995

Variables	Absolute Value Coefficient
Distance2LavaFlow1250	-.173
Distance2ExistingMF1	-.137
Distance2Airports	.123
Distance2SewerService	-.106
Distance2VolcanoHazard	.079
Intersections1miDensity	-.078
Distance2Schools	.073
Distance2WaterService	.073
Proximity10ftParcelDensity	.064
Proximity1000ftParcelDensity	-.059
Distance2LavaFlow1790	.051
Slope	-.049