# Countrywide Automated Property Evaluation System - CAPES

**Ingemar A.E. Hulthage and Iain Stobie**

Countrywide Home Loans, Artificial Intelligence Division, 55 S. Lake Avenue, Pasadena, California 91109
Email <firstname>_<lastname>@Countrywide com

## Abstract

The purpose of CAPES is to estimate the market value of residential properties in order to assess the collateral on Countrywide mortgage loans. CAPES estimates market value by comparison of the subject property to other similar nearby properties, for which recent sales information is available. In some cases price indices describing the change in property values over time are also used. In addition to the estimated market value, CAPES produces a measure of the uncertainty in the result. It uses several models, including heuristics derived from company-specific business rules, and accesses both commercial and proprietary property databases. Its accuracy has been validated extensively on batches of properties by comparing its results to known sales prices. It is integrated with Countrywide's underwriting expert system and is currently being used by over thirty departments on a daily basis.

## Problem Description

Countrywide is the nation's leading independent residential mortgage lender, currently funding over 50 thousand new loans a month and servicing mortgages for more than 1.7 million homes. Mortgage banking consists of three major activities: originating mortgages from borrowers or brokers, or purchasing mortgages from banks or credit unions; selling mortgages (usually as part of a large pool of loans) to secondary market investors; and servicing mortgages on behalf of the investors (collecting monthly mortgage payments, addressing requests to refinance or cancel mortgage insurance, and handling problems such as late payments and foreclosures).

At many stages in these processes an assessment of the market value of the property associated with a mortgage is of major importance. Market value is defined by federal financial institutions as:

> "... the most probable price a property should bring in a competitive and open market under all conditions requisite to a fair sale, the buyer and seller each acting prudently, knowledgeably, and assuming the price is not affected by undue stimulus."

Appraisals are the primary means for estimating the market value of properties. Traditionally conducted by certified appraisers using the market comparable approach, a few similar neighboring properties (comparables or "comps" for short) that have sold recently are used to estimate a property's market value. (The current discussion only concerns the valuation of residential single-unit homes; different appraisal techniques apply to investment, commercial and industrial properties.) Ideally the selected comparables should be identical to the subject property (model matches). However, such comparables are rarely available and therefore the appraiser adjusts for the differences, guided by accepted principles of appraising. In addition, the subject property is physically inspected and neighborhood trends evaluated. An appraisal thus formalizes the process knowledgeable homebuyers use when purchasing a home.

Accurate property valuation is difficult because no truly "objective value" exists. The purchase of a home is not a purely rational process. Both buyer and seller have limited information and are subject to personal taste and needs. For example, the number of bedrooms is of greater importance to homebuyers with children than to those without children. The price is also dependent on the negotiation skills of the parties involved. These and other factors cause the value of a property to have an inherent spread, in that repeated sales of the property under identical conditions would yield a distribution of prices. The standard deviation of this distribution has been estimated to be 5%-7% of the average value (Case & Shiller 1987).

The impact of property characteristics on value is also hard to quantify. Although values can be calculated by estimating construction costs plus land value (the "cost approach"), this method is unreliable and not often used. The most important characteristics of a transaction – the time frame of the sale and location of the property -- are so hard to quantify that they are factored out of the process entirely by considering only "recent neighboring" comps.

Appraisals can create a bottleneck in originating a mortgage and their manual nature makes them inappropriate for bulk transactions such as estimating the total market value of a portfolio of loans and data mining (e.g., to market products to borrowers with equity in their home).

## Objectives of Automated Property Valuation

Although statisticians have studied the problem of automated property valuation for decades, only in the last few years has the advent of commercial property databases made implementation feasible.

The major goal is price reduction and improved speed, thereby creating a more efficient origination process for both borrower and lender. Automation also offers objectivity (by minimizing the individual taste of a particular buyer/seller or pressures on the appraiser) and consistency (by minimizing individual differences between appraisers).

However, there are issues and aspects of the appraisal process that automated systems will probably never be able to adequately address. Specifically:

- Property databases, usually derived from public records, cover only about half the country and can be incomplete and inaccurate.
- Available databases have no reliable information about the condition of properties. The fact that a property has been constructed, improved, or even destroyed may not be reflected in available databases.
- Other data (e.g., view) may be subjective or imprecise.
- Heterogeneous neighborhoods and unusual properties pose additional problems.

For these and other reasons it is neither appropriate nor desirable to eliminate human involvement from appraisals. Instead, the CAPES project provides an alternative property valuation method that can be used to complement manual appraisals or in situations where the advantage of a human appraisal is less significant (Eckert et al. 1993). Countrywide has a variety of property valuation needs and requires a system that supports a wide range of uses, from an interactive appraiser's assistant that can be tailored using expert knowledge of the neighborhood or subject property, to a fully automated expert system that can be used by novices or by a batch process. Depending on the other risk factors on the loan, a physical property inspection may still be required to double-check the accuracy of the data.

### Approaches to Automated Property Valuation

All approaches follow the same basic strategy.

- Define a set of significant residential property features
- Collect information on these features for the subject and nearby properties
- Develop a model to estimate the expected sales price of the subject property based on the selected features
- Apply the model to the subject to compute a value, and preferably an estimate of confidence in the value

The approaches differ in the features defined as significant and the type of model used. At one extreme are "Home Price Index" models, using only sales prices and dates to calculate what amounts to the "average" price change in an area (zip code or county) over time (Case et al. 1991, Shiller 1991). Although this provides a useful "macro" model, it does not take into account improvement or deterioration of the subject over time and its accuracy decreases the older the previous sales date. Most models use additional property features, including: physical characteristics such as living and lot area, number of bedrooms and bathrooms, age, and presence of pool, garage, etc.; location identification such as street address, latitude and longitude, and census tract; and qualitative information such as view. Models that are based on property features are sometimes called hedonic. The property information usually comes from county assessors and recorders offices, cleaned up and repackaged by commercial vendors in electronically accessible databases. Another source is proprietary property and appraisal databases maintained by large lenders such as Countrywide.

Dimensions and volume of information distinguish human from automated valuation. Human appraisers (and homebuyers) perceive broad and detailed information, including such hard to quantify features as view, condition, "flow" and "appeal," from the visual inspection of only a few comparables. In contrast, automated systems can review hundreds of comparables but have access to only specific, limited information on each.

**Linear Regression** has long been applied to this problem, modeling price as a linear function of the property features and thereby automatically determining what adjustments should be made for differences in features. However, straightforward regression fails to consistently yield accurate results, as has been repeatedly reported in the literature (Murphy 1988). The set of properties that are available for the regression analysis often is too heterogeneous (Newsome & Zietz 1992) and the technique too sensitive to correlation between variables (for example, the numbers of bedrooms and bathrooms).

**Non-Linear Regression** and other variations on basic linear regression have been proposed (Do & Grudnitski 1992; Knight et al. 1993), but they involve additional assumptions or may require more input (Weirick & Ingram 1990).

**Neural Networks** are used by some systems (Jost et al. 1994). However, the unfeasibility of learning in real-time necessitates off-line learning, resulting in a macro model of neighborhoods, not individual properties.

**Rules-based methods** are a natural choice for this problem, because human appraisals are based on established principles and professional guidelines, such as the "Uniform Standards of Professional Appraisal Practice" (USPAP). Although an element of judgment is often involved, these rules can be applied in a principled manner that is well described in several texts (Betts & Ely 1994). Additional expertise can be drawn from interviewing professional appraisers. A final advantage of these methods over "black box" regression and Neural Network techniques, is that the declarative representation of rules can be used to explain why the result was obtained.

## Application Description

Countrywide's overall computing infrastructure is a decentralized collection of personal computers and servers connected by local area and wide area networks.

CAPES is a client / server application implemented in MS Visual C++ and based on the Distributed Common Object Model (DCOM) (Grimes 1997). DCOM simplifies the implementation of client / server applications by allowing the creation of COM objects remotely on another machine. This architecture makes it possible to create a variety of clients served by the same server. In addition to the client

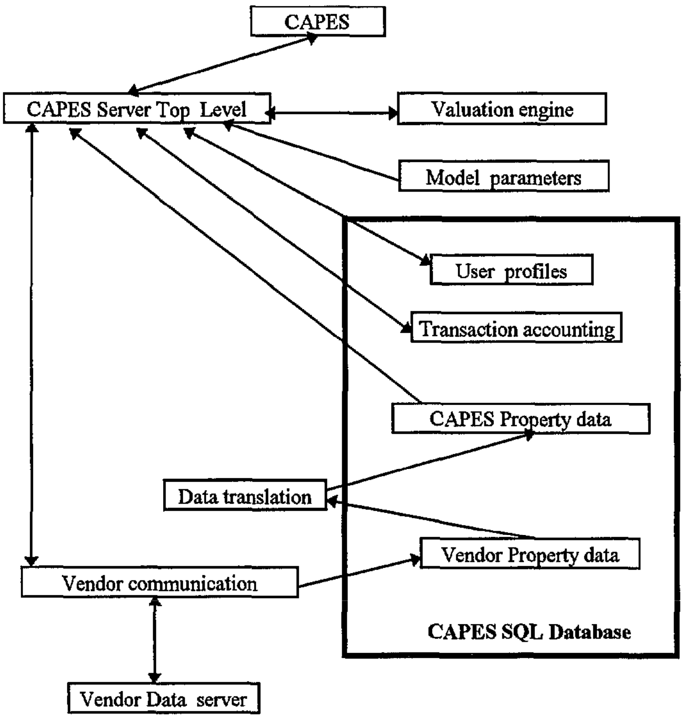


**Figure 1: CAPES is a client/server system, where the server handles all communication with internal and external data sources.**

described in this paper, there is a client that interfaces with CLUES (Talebzadeh et al. 1995) and another that makes CAPES available on the Internet/Intranet.

The CAPES server application runs as a service on a Windows NT 4.0 server and handles all users. The client application runs on the user machine, which may be running Windows NT 4.0 or Windows 95 (with a MS patch to support DCOM). The connection is normally made over a WAN, but can also be made via RAS.

The server application is connected via a telnet connection over a leased phone line to DataQuick, our main present property data supplier. All property data and most results from CAPES runs (including usage statistics, error logs, etc.) are permanently stored in an SQL database. The client applications have no direct connection to the database or to the data suppliers.

In general CAPES is designed from the ground up in an object oriented style, using derived and template classes, virtual functions, etc. The data representations are defined by classes. Much of the computation is performed by creating objects of specific classes, which encapsulate the input data, the computation or reasoning, and the result.

CAPES also uses the C++ Standard Template Library (STL) (Stepanov & Lee 1995), a general-purpose library of generic data structures (such as lists and maps) and operations (such as insertion of elements and sorting). STL makes programming more productive by handling memory management and providing a framework for decomposing many programming problems.

**CAPES client**

The client application receives and displays all information on the subject property, control information from the user, and qualitative and quantitative information, including the estimated value with uncertainty and status information.

The client user interface is implemented using MS Foundation Classes (MFC) and third party OCX controls. The client is a Multi Document Interface (MDI), where each subject property is assigned a separate window (Figure 2) consisting of five tabbed child windows (Subject, Results, Comp Selection, Market Analyses, and Messages). Each tab provides specific user interface features.

CAPES allows a large degree of flexibility in the level of control and expertise required from the user. It begins by performing an automatic evaluation that requires minimal input and direction from the user. This mode is convenient and suitable for users without in-depth knowledge of appraisal methodology. Then, depending on the estimated uncertainty of the results, or for other reasons, users can:

- Edit the information on the subject property and control the selection of comps in detail
- Evaluate the result based on detailed quantitative and qualitative output
- Iteratively repeat this process and compare the results

To start an evaluation, users must uniquely identify the subject property, usually by specifying the street number, name, and zip code, as shown in the upper part of Figure 2. When the Find Subject button is selected CAPES searches the data source for matching property data and, if found, displays it on two lines, as seen in the middle of Figure 2. The lower line is the data retrieved from the data source while the upper line is the data actually used in the evaluation. The data used by CAPES is, by default, equal to the data retrieved, but the user can edit any incorrect information on the upper line. In the figure the user has changed the number of bathrooms from 1.70 (i.e. 1¾) to 2.00. These lines scroll horizontally and contain additional characteristics.

When satisfied that the right property has been found and that its characteristics are as accurate as possible, the user starts the evaluation by selecting the Compute button. CAPES attempts to find comparable properties and performs the steps of the automated evaluation (described below). The result is displayed on the bottom line and includes: the estimated value; a measure of uncertainty, expressed as a percentage of the value; the number of comps

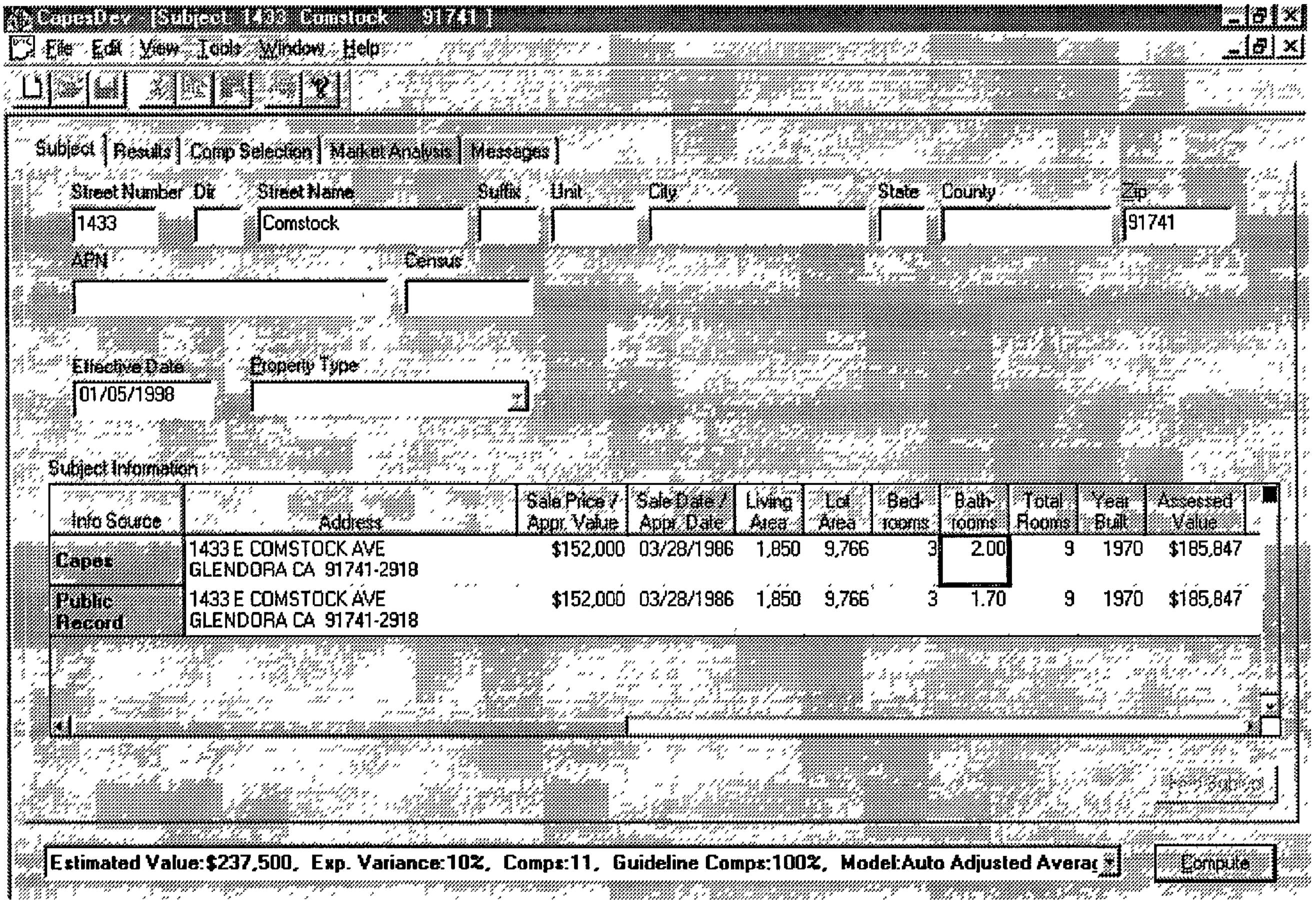


**Figure 2: An automatic CAPES appraisal only requires minimal property identification input**

used; the percentage of the comps that meet guidelines defined by the major funding agencies; and the model used to make the estimate. (The guidelines for comps apply to traditional appraisals, which may include as few as two comps. When CAPES uses many more comps, they are not all required to meet these guidelines.)

Expert users with detailed knowledge of the appraising process, or perhaps of the subject property or area, may further customize the evaluation using the functionality implemented on the remaining tabs.

- The Comp Selection tab initially displays the filters selected by CAPES and a collection of statistics on the current comp set. A filter is a parameterized operator to subset collections of properties. Many simply choose an interval for a feature, e.g. 2-3 bedrooms.

  Using the selected filters and statistics as a guide, users may modify (completely or in part) the set of filters that determine the selection of comps. This is the most important control action available to users.
- The Market Analysis tab displays the comp set in a tabular format similar to the display of subject characteristics shown in Figure 2. Expert users may eliminate any comp from the comp set.
- The Results tab displays the results of all applied valuation models. It also contains other qualitative and quantitative information to help expert users judge the reliability of the estimated value. For example, if the comp set is biased because most comps are either more or less valuable than the subject, a warning message displays on the Results tab. In most such cases no value is displayed on the bottom line.
- The Messages tab displays error and other miscellaneous messages that may be of interest to expert users. All users are alerted to serious failures, such as the failure to find subject or comp data, through pop-up windows.

A batch facility utilizing the automatic mode is also available. Users insert property identification data in a template table stored in a file. The file is read by CAPES and the properties evaluated one by one. When the evaluation is complete, results are reported in a file merging the original input with a number of columns holding the corresponding results. Once started this batch mode runs unattended, making it easy to evaluate portfolios of properties.

## CAPES server

The server top level provides the methods available to clients and controls the interaction with clients. It also coordinates finding subject and comp information and invoking the valuation engine.

## CAPES Valuation Engine

The valuation engine uses available information on the subject property and a relatively large set of property sales. If possible, more than 150 comparable sales are retrieved from a commercial data vendor. The only restrictions on this initial set are that they must be fairly recent full value sales within a reasonable distance of the subject property.

To achieve a valuation with a low uncertainty it is usually necessary to reduce the set of comps by applying one or more filters. CAPES automatically selects a set of filters.

Given a set of comps, the value of the subject is estimated by several models. These models include regression and other models based on statistical analysis, as well as an adjustment model which seeks to follow the practices of appraisers. Regional home price indices are also used when historic sales price and date information is available. These models are described in some detail in the next section.

An important issue is that information is often missing for the subject or one or more comps. Missing information is treated as a special value, rather than as zero or some other default value, and the models are designed to accommodate this.

The result of each model is checked against a set of constraints to ensure that a lack of information (or some other reason) does not produce an invalid value. The uncertainty in each model result is estimated as the standard deviation in the subject property value. The model with the lowest estimated uncertainty that also satisfies the checks is returned as the estimated value. If no model satisfies the constraints, no value is returned to the user.

### Model parameters

Some details of the valuation models are controlled by model parameters defined in the code or the database. The most important model parameters are adjustments for property features. An adjustment parameter assigns a dollar value to a feature (e.g., a bathroom might be worth $1,000) based on expert knowledge. Adjustments are collected in tables and are differentiated by criteria such as price tiers and property types.

### User profiles

Access to CAPES is controlled by the user's network logon. A user table in the database grants privileges ranging from no access to full access, and also contains contact and cost accounting information.

### Transaction accounting

To monitor the use of CAPES and any errors that may occur, and to allow CAPES results to be reconstructed, each transaction, along with all data and results, is logged in the database.

### CAPES property data and translation

CAPES stores the raw property data from vendors and internal sources verbatim and translates from these formats into a single internal format that is used by the valuation engine A translation module implements the semantics required to interpret the various representations and conventions of these data sources.

### Vendor communication

Communication with the data vendor is via a telnet connection over a leased phone line.

## Reasoning about property evaluation

The main inference techniques used in CAPES are described in this section.

### Statistical models

The median or average value of the filtered comp set is sometimes the best model of value. Statistics based on other sets, such as the neighborhood or the nearest neighbors, are also of interest.

### Regression model

The regression model performs a linear regression of the sales prices in the set of comparable properties. Property characteristics are used as independent parameters and the regression function is used to model the value of the subject property. It is implemented with the singular value decomposition method, which is robust against missing and redundant data. (Press et al. 1992, Jefferys 1980,1981; Lybanon 1984)

### Adjustment model

The adjustment model seeks to follow the practices of appraisers. The known sales prices of a set of comparables are adjusted based on differences in the characteristics of the subject and each comparable property. The average of the adjusted sales prices is the value estimated by the adjustment model.

### Home price indices

Price indices are constructed by reviewing repeated sales of properties and calculating an index value for a point in time. Given the sales price of a property at a specific time, the value at any other time (within the period for which the index is valid) can be computed by multiplying the original sales price by the ratio of the index values. To improve the accuracy a collection of indices is used, with different indices for different property types, price tiers, and areas.

### Best first search

CAPES differs from other property valuation systems by using heuristic search as the core of the evaluation engine.

The main principle of appraising is to select a set of "comparable" properties that have recently sold in so-called "arms length" transactions. If the concept of comparable property was well defined, and if sufficient information to identify comparable properties was always available, there would be no need to search. The only thing needed would be a set of rules, or constraints, to classify comps as suitable or unsuitable for a specific appraisal.

However, experience shows that when a fixed set of criteria are applied, in many cases either a too small or an unnecessarily large set of comparables is selected. (A comp set can be too large in the sense that a more accurate evaluation can be obtained from a smaller, more carefully selected set of comps.) Further reducing the usefulness of fixed criteria is the uncertainty of the information on the subject and comparable properties. Several characteristics, such as the condition of the property, are not captured in the public records, while other information (on the subject or one or more comps) may be missing or simply wrong.

This situation suggests searching for a set of comparables that optimizes the accuracy of an appraisal based on that set. If an initial set of N recent transactions in the subject neighborhood is chosen, the search space of all subsets is of the order of $2^N$. For example, recent transactions in the neighborhood could mean all transactions in the last 12 months within a radius of one half mile from the subject. In a typical suburban area this would often include more than 100 potential comparables. The corresponding search space of subsets is too large to search exhaustively; fortunately, as explained below, an exhaustive search is not necessary or desirable.

The goal of the search is to minimize the error in the estimated value of the subject property. Because the true error is not known, the search must be guided by an estimate of the error, such as the standard deviation of the sales prices of the comp set. However, this evaluation function tends toward over-fitting. That is, the search finds a set of properties that happen to have a small range of sales prices (and therefore generate a low error estimate) but aren't otherwise particularly suitable as comps.

The operators in the search select a subset of a given set and are therefore called filters. By using filters that systematically generate all the subsets of the initial set, an exhaustive search has an overwhelming risk of over-fitting. To reduce the complexity and minimize the problem of over-fitting we only use search operators that are heuristically well motivated.

A heuristic property filter excludes properties from the comp set based on rules known by appraisers to be useful. Such rules generally focus on property characteristics. For example, assume the subject has three bedrooms and the comp set contains properties with either three or four bedrooms. Heuristic filtering based on the number of bedrooms would yield only two nodes: the original set and the subset with exactly three bedrooms. Other filters (based on the number of bathrooms, living area, etc.) typically generate additional nodes. However, the number of nodes is reduced by dependencies among the operators. For example, applying a filter limiting the living area to ± 10% difference from the subject frequently yields a subset with all remaining comparables having the same number of bedrooms. In this case, applying the bedroom filter doesn't produce a new node.

Heuristic filtering reduces the search space dramatically. If there are 10 commuting filters, the search space is now at most on the order of $2^{10}$ nodes. The number of nodes generated is also limited by the Best-First search, which is well described elsewhere (Rich 1983). Furthermore, because of the inherent spread in a property's value, nothing is gained by continuing the search when a node with an error estimate below some level (e.g. 5%) is found. Finally, experimentation shows that the search can be limited to a fairly small number of nodes (e.g., 100) without increasing the average real error. As explained above, the risk of over-fitting out weighs the advantage of generating more nodes.

### Outlier detection and removal

The inherent spread of property prices, and the fact that the properties in a comp set are not identical, implies that the prices in a comp set is a statistical distribution. It's reasonable to assume that this is approximately a normal distribution. However, one frequently observes sales prices that are many standard deviations away from the average. There are many possible reasons for anomalous sales prices, such as a distressed sale, a sale within a family, or extraordinary features that reduce or increase the property's value. Such data points are often called outliers (Barnett & Lewis 1995). CAPES uses statistical criteria to remove comps that are outliers in price and other selected characteristics.

### Uncertainty model

An important part of CAPES is the attempt to estimate the standard deviation in the estimated value. This modeling of uncertainty is based on statistical principles and including contributions from different sources of uncertainty, such as differences between subject and comps or missing information.

### Incomplete information

The information on the subject and comparable properties is frequently incomplete. Almost any attribute can be randomly or systematically missing. Examples of systematically missing information are the availability of total number of rooms and number of bedrooms. In some areas both attributes are provided, while in other areas only one or neither are provided. CAPES doesn't presently utilize knowledge about such systematic differences; instead, all missing values are represented as special values (distinct from valid values) and treated appropriately. For example if the number of bedrooms is missing for the subject, no adjustments are made for the comps based on number of bedrooms. This approach has the advantage of being robust against randomly missing information while avoiding the overhead of maintaining a knowledge base of regional differences in data availability.

### Constraint reasoning

A critical part of the valuation process is critique of the model results. There are many rules that can be applied to a model's results to determine weather the result is reliable or if it is more or less accurate than the results of other models. These rules are either heuristic or statistical, and are expressed in CAPES as constraints, using a simple constraint propagation system (Tsang 1993) which finds the subset of model valuations satisfying all constraints. If more than one model remains, the one with the lowest uncertainty estimate is chosen. If no model remain, a failure is reported.

## Application Development and Deployment

Since its creation in late 1991 the Countrywide AI department has developed a number of applications (Stobie 1996). A study of selected automated property valuation systems was done in 1995. No system allowed the degree of control that our expert users require or allowed us to use company-specific expertise, rules of business, and proprietary appraisal databases. After deciding that no commercial system met Countrywide's broader needs the CAPES project was initiated in February 1996. By July 1996 a prototype demonstrating the feasibility of the project had been developed and received the support of senior management. It focused on the core AI problem of computing a property's value from a set of comps, using a simple user interface and commercially available data.

A full-scale development effort took place to implement a client / server architecture, access remote and internal databases, and create a user-interface powerful enough to support the functions required of an appraiser's assistant. In parallel, the AI and statistical models were continually refined based on ongoing knowledge acquisition and experience of using the system. In May 1997 the roll out of CAPES began and has since been deployed to a growing number of departments. A User's Manual and periodic training sessions adequately prepare users for CAPES, while the developers themselves support the product via telephone and e-mail. In October 1997 CAPES was formally demonstrated at the Mortgage Banking Association conference in New York city.

Object-oriented design and analysis together with rapid prototyping was used to develop CAPES. Each time a change to the core valuation engine is proposed, it is implemented as quickly as possible and a batch test run to determine the effect of the change on the accuracy and coverage of the system.

- Accuracy is measured by how closely the estimated values compare to the known sales prices, according to average error, average absolute error, maximum error, and distribution of errors. Comparison to the subject sales price is not a perfect benchmark (because of the occasional forced sale or other non-arms-length transaction), but we believe it is the best available.
- Coverage is measured by how often a subject property is found and a final result produced.

The fact that all CAPES runs are archived in a database allows re-running batch tests of thousands of properties in minutes, generating statistics on the results, and judging whether the changes have been implemented correctly and have improved performance. We are also benchmarking CAPES against other valuation systems, as opportunities arise.

## Application Use and Payoff

CAPES is used by over 30 departments at Countrywide, many on a daily basis. The marginal cost of a CAPES valuation (costing only a few dollars and taking only a few minutes) makes it practical for a wide variety of applications.

- CAPES is used to analyze the collateral value of a portfolio of loans purchased from a smaller bank as part of the "due diligence" process. Because CAPES is statistically nearly unbiased, its average signed error is close to zero—the over-estimates cancel out the under-estimates. By randomly sampling a subset of the portfolio of loans in regions covered by our property databases and running the properties through CAPES, a quick and reliable total value of the portfolio can be obtained. Since these properties have already been appraised, this also facilitates a focused quality control procedure: loans where the appraised value differs greatly from the CAPES value are carefully reviewed.
- This use of CAPES to review appraisals is also applied by the Quality Control, Internal Audit, and other departments.
- The Foreclosure department uses CAPES to estimate the potential loss (the difference between the estimated property value and the balance of the loan) incurred when a loan defaults.
- Once a loan has been foreclosed and the property belongs to Countrywide, the Real Estate Management department uses CAPES as an aid to determine a reasonable sales price.

The challenge to using CAPES for loan origination is that most investors (to whom we deliver the loans) have not yet approved automated valuation as an acceptable appraisal method.

Integrating CAPES (automated property valuations) and CLUES (automated credit underwriting) potentially increases the accuracy of both systems. For example, CLUES verifies that the comparables used by the appraiser are within a reasonable distance of the subject, that they are not all more or less expensive than the subject, and that the adjustments are not excessively high. CAPES enhances CLUES by verifying the accuracy of the subject and comps used in the appraisal, and providing a second, independent estimate of value.

Speed is of the essence in today's competitive mortgage lending business and automation is a major success factor. With CLUES analyzing the credit history and financial ability of the borrowers to repay the mortgage and CAPES evaluating the collateral, AI at Countrywide provides a quick, inexpensive, and increasingly accurate "streamlined" approval of low-risk loans. Countrywide is considering offering this capability directly to customers, for example over the Internet.

## Maintenance and Planned Enhancements

The CAPES development team continues to maintain and enhance the system. Weekly meetings with experts address issues based on feedback from the large user base and new industry directions. The near monthly release of enhanced versions of the software is greatly facilitated by the use of object-oriented design, cleanly separated modules, and declarative rules.

Several enhancements are planned, including more sophisticated spatial reasoning using location of the subject and maps. For example, it is currently quite possible that nearby properties are chosen as comps when in fact they are from an adjacent, but very different neighborhood.

Additional data sources are required to achieve higher coverage and more detailed and accurate data. To permit valuations in less than a second, it is imperative to install these databases locally at Countrywide, rather than accessing them remotely via a leased line. This opens the door to using CAPES for data mining applications, such as identifying candidates from Countrywide's portfolio of 1.7 million loans that may benefit from refinance or are in danger of foreclosure. Data mining techniques could also automate some of the knowledge acquisition, such as empirically determining the value for adjustments. For example, the value of a swimming pool in a specific neighborhood (perhaps as a function of other features) can probably be derived by comparing a number of properties in that neighborhood

## References


Barnett, V. and Lewis, T. 1995, *Outliers in statistical Data*, John Wiley & Sons.

Betts, R.M. and Ely, S.J. 1994. *Basic Real Estate Appraisal.* 3rd edition, Prentice-Hall.

Case, K.E., Bradford, Pollakowski, H., and Wacter, S., 1991. On choosing among house price index methodologies. *AREUEA Journal*, 1991:19, 286-307.

Case, K.E., and Shiller, R..J. 1987, Prices of Single Family Homes since 1970: New indexes for Four Cities. *New England Economic Review*, 1987, 45-56.

Do, H and Grudnitski. 1992. A Neural Network Approach to Residential Property Appraisal. *The Real Estate Appraiser* December 1992:38-45.

Eckert, J.K., O'Connor, P.M. and Chamberlain, C. Computer-Assisted Real Estate Appraisal: A California Savings and Loan Case Study. *The Appraisal Journal*, October 1993:524-532.

Grimes, R., 1997. *Professional DCOM programming*, Wrox Press Ltd.

Jefferys, W.H. 1980. On the method of least squares. 1980,1981 *The Astronomical Journal* 85(2):177-181 and 86(1):149-155. Errata 95(4).

Jost, A., Nelson J., Gophinatan, K., Smith C., Real estate appraisal using predictive modeling. U.S. Patent 5,361,201, Nov 1, 1994.

Knight, J.R., Hill R.C. and Sirmans, C.F. 1993. Stein Rule Estimation in Real Estate Appraisal. *The Appraisal Journal*, October 1993:539-544.

Lybanon. M. 1984. A better least-square method when both variables have uncertainties. *Am. Journal of Phys.* 52(1):22-26.

Murphy L.T. III. 1989. Determining the Appropriate Equation in Multiple Regression Analysis. *The Appraisal Journal*, October 1989:498-517.

Newsome, B. A. and Zietz, J. 1992. Adjusting Comparable Sales Using Multiple Regression Analysis – The Need for Segmentation. *The Appraisal J.*, January 1992: 129-135.

Press W.H. Teukolsky S.A., Vetterling W.T. and Flannery, B.P. 1992. *Numerical Recipes in FORTRAN.* 2nd edition, Cambridge University Press.

Rich, E., 1983. *Artificial Intelligence*, pp. 78, McGraw-Hill.

Shiller, R..J. 1991. Arithmetic repeat sales price estimators. *Journal of Housing Economics*, 1991:1, 110-126.

Stepanov, A.A. and Lee, M., 1995. The Standard Template Library, Hewlett-Packard Technical Report, HPL-94-34, revised 1995.

Stobie, I., Artificial Intelligence at Countrywide, Proceedings of Wescon/96:468-472.

Talebzadeh, H., Mandutianu, S., Winner, C. F., Countrywide Loan Underwriting Expert System, AI Magazine, Volume 16, No. 1, 51-64, Spring 1995

Tsang, E., 1993. *Foundations of Constraint Satisfaction*, Academic Press.

Weirick W.N. and Ingram F.J. 1990. Functional Form Choice in Applied Real Estate Analysis. *The Appraisal Journal*, January 1990:57-73.