Geologically constrained electrofacies classification of fluvial deposits: An example from the Cretaceous Mesaverde Group, Uinta and Piceance Basins

Daniel B. Allen and Matthew J. Pranter

ABSTRACT
Statistical classification methods consisting of the $k$-nearest neighbor algorithm ($k$-NN), a probabilistic clustering procedure (PCP), and a novel method that incorporates outcrop-based thickness criteria through the use of well log indicator flags are evaluated for their ability to distinguish fluvial architectural elements of the upper Mesaverde Group of the Piceance and Uinta Basins as distinct electrofacies classes. Data used in training and testing of the classification methods come from paired cores and well logs consisting of 1626 wireline log curve samples each associated with a known architectural element classification as determined from detailed sedimentologic analysis of cores ($N = 9$). Thickness criteria are derived from outcrop-based architectural element measurements of the upper Mesaverde Group. Through an approach that integrates select classifier results with thickness criteria, an overall accuracy (number of correctly predicted samples/total testing samples) of 83.6% was achieved for a four-class fluvial architectural element realization. Architectural elements were predicted with user’s accuracies (accuracy of an individual class) of 0.891, 0.376, 0.735, and 0.985 for the floodplain, crevasse splay, single-story channel body, and multistory channel body classes, respectively. Without the additional refinement by incorporation of thickness criteria, the $k$-NN and PCP classifiers produced similar results. In both the $k$-NN and PCP techniques, the combination of gamma ray and bulk density wireline log curves proved to be the most useful assemblage tested.

ACKNOWLEDGMENTS
Funding was provided by the industry sponsors of the Reservoir Characterization and Modeling Laboratory and the Williams Fork Consortium: Anadarko, Chevron, ExxonMobil, Occidental (Oxy), QEP Resources, Suncor Energy, and WPX Energy. Additional funding was provided through the AAPG Foundation Grants-in-Aid (Norman H. Foster Memorial Grant) and the SEPM Endowment Grant. Software was provided by IHS (Petra); Fugro-Jason (Powerlog 3.3 and FaciesID); and eGAMLS Inc.
INTRODUCTION

The classification of architectural elements is a vital task in the investigation of fluvial depositional systems and reservoirs. This task is commonly conducted through the study of core samples or through observations made from outcrop. Numerous such reservoir-scale studies (Ellison, 2004; Cole and Cumella, 2005; German, 2006; Panjaitan, 2006; Pranter et al., 2007, 2009; Cole, 2008; White et al., 2008; Pranter and Sommer, 2011) of the fluvial deposits of the upper Mesaverde Group of the Piceance and Uinta Basin province of western Colorado and eastern Utah have recorded detailed sedimentologic analysis, dimensional characteristics (e.g., width, thickness, and width-to-thickness ratio), and paleocurrent data to describe and classify the different fluvial architectural element types common to these deposits. These efforts to classify the fluvial architectural elements and record their distributions have resulted in a better understanding of the fluvial depositional system of the upper Mesaverde Group. Other studies have, at least in part, if not fully, relied on the manual interpretation of fluvial architectural elements in well logs through commonly recognized well log curve motifs (Rider, 2002) to assist in the investigation of the sequence stratigraphy of the upper Mesaverde Group (Shaak, 2010) and the static connectivity of the reservoir-quality sandstones that compose this interval (Hewlett, 2010; Sloan, 2012; Pranter et al., 2014).

Though extensive study has been conducted through the analysis of core and outcrop exposures as well as the manual interpretation of fluvial deposits from well logs, these methods have their constraints. Most wells have well logs but are not cored, whereas outcrops suitable for study can be geographically limited. Though comparatively more economical than coring wells to aid in interpretation of fluvial deposits, manual interpretation of these deposits using well logs can be a tedious, time-consuming, and subjective task for even the most experienced of well log analysts and geologists.

As a response to these constraints, this study explores an alternative approach to the classification of fluvial architectural elements through the application of two statistical classification methods: (1) the k-nearest neighbor algorithm (k-NN) and (2) a probabilistic clustering procedure (PCP). In addition, a third approach is used to refine the results of the classifiers, which incorporates outcrop-based, architectural element thickness criteria. These methods are evaluated for their ability to extrapolate the fluvial architectural elements as distinct electrofacies classes from training wells (wells with core and log data used to train the classification method) to testing wells through the comparison of wireline log curve measurements. Data used in the training and testing of the classifiers come from (N = 1626) wireline log curve samples (values) associated with known architectural element
classifications as determined from detailed sedimentologic analysis of cored wells (N = 9) located throughout the study area (Figure 1). The results of this study provide a methodology for making interpretations of the fluvial architectural elements using well logs that is carried out in a cost-effective, timely, objective, and reproducible manner.

The discrimination of depositional facies from well logs dates to the mid-1950s, when workers at Shell used spontaneous potential well log curve shapes to distinguish depositional facies of the modern Mississippi Delta (Serra, 1989). Studies in the 1970s (Serra and Sulpice, 1975; Rider and Laurier, 1979; Serra and Abbott, 1982) attributed characteristic shapes of additional well log curves such as gamma ray (GR), bulk density (RHOB), neutron porosity (NPHI), and dipmeter logs to depositional facies (Sullivan et al., 2003). With the ever increasing need for timely and cost-effective methods of facies classification, the decades since the 1980s have seen a growth in the application of various statistical approaches to automated electrofacies prediction. The term “electrofacies” was used by Serra and Abbott (1982, p. 117) and is defined as the set of log responses that characterizes a bed and permits it to be distinguished from others, a task typically carried out using simplistic graphical techniques (Doveton and Prensky, 1992). Wolff and Pelissier-Combescure (1982) were early practitioners of the multivariate statistical approach to automated facies when they used principle component and cluster analysis to estimate the occurrence of lithofacies. The use of

Figure 1. Map of the study area in the Piceance and Uinta Basins of northwestern Colorado (CO) and eastern Utah (UT), respectively. Cored wells used in training and testing classifier models are shown as well as noncored wells, which were used for demonstrating the prediction and mapping of electrofacies classes across the region. Lat = latitude; Long = longitude.
multivariate statistics continued with the implementation of discriminant function analysis by Busch et al.
(1987) and Delfiner et al. (1987) to estimate the occurrence of lithofacies (Dubois et al., 2007). Perhaps the most popular method of electrofacies classification in recent years (e.g., Kapur et al., 1998; Grotsch and Mercadier, 1999; Saggaf and Nebrija, 2000; Russell et al., 2002) has been the nonmultivariate statistical approach of the artificial neural network, which Dubois et al. (2007) showed to have an advantage in its ability to correctly predict electrofacies classes when compared with other commonly used classification methods. A summary of commonly performed statistical approaches to electrofacies classification can be found in Doveton (1994). As in this study, many of the previously mentioned studies use core-defined depositional facies to provide ground truth to the predictive capabilities of the classification methods to facilitate acceptance of the results among other geoscientists.

This study examines the statistical approaches of the k-NN algorithm and a PCP. The k-NN algorithm (Cover and Hart, 1967) is attractively noted as being one of the simplest and most intuitive classification methods (Bremner et al., 2005; Hall et al., 2008). Unlike complex artificial neural network classifiers, this classification method is a simple look-alike contest where unknown objects are matched according to similar objects with known classes (Dubois et al., 2007). Though simplistic, in a study comparing the ability of four commonly used classifiers’ abilities to predict selected lithofacies of the Permian Council Grove Group in the Hugoton and Panoma fields in southwest Kansas and northwest Oklahoma, Dubois et al. (2007) found the k-NN classifier to perform comparatively well. The k-NN classifier has also been applied to distinguishing between ground cover classes as captured by satellite imagery and is used as a tool in internet search engine functions (McRoberts et al., 2002; Haapanen et al., 2004; Beaudoin et al., 2005; Yeung et al., 2008). The PCP is a maximum likelihood model-based neural system (but not an artificial neural network) (Eslinger and Boyle, 2011). It is the main clustering engine within the Geologic Analysis via Maximum Likelihood System™ software. The PCP has been previously employed in past studies in the discrimination of lithofacies of the Mississippian Barnett Shale at Newark East field, Fort Worth Basin, Texas (Vallejo, 2010), and has been applied to the estimation of petrophysical properties (Eslinger and Boyle, 2011).

TECTORIC AND STRATIGRAPHIC SETTING

The Campanian-age Uinta and Piceance Basins are located in northeastern Utah and northwestern Colorado, respectively (Figure 1). The Uinta Basin is asymmetrical with a west–northwest-trending axis. It is roughly 120 mi (193 km) in length and nearly 100 mi (161 km) wide. The basin is bounded by the Uinta uplift to the north, the Wasatch Plateau to the west, the Rafael and Uncompahgre uplifts to the south, and the Douglas Creek arch to the west (Spencer, 1995; Stancel et al., 2008). The Piceance Basin is asymmetrical with a northwest–southeast-trending axis. It is 100 mi (161 km) in length and 40–50 mi (64–80 km) wide (Spencer, 1995). The basin is bounded to the northwest by the Uinta uplift, to the north by the Axial arch, to the east by the White River uplift, to the southeast by the Elk Mountains and Sawatch uplift, to the south by the Gunnison uplift and San Juan volcanic field, and to the southwest by the Uncompahgre uplift and is separated from the Uinta Basin by the Douglas Creek arch to the west (Johnson, 1989). During the Cretaceous, both the Piceance and Uinta Basins were part of the much larger Cretaceous Rocky Mountain foreland basin that formed as a result of thrust loading along the Sevier orogenic belt, which lies on the western margin of the Uinta Basin. From Late Cretaceous through the Eocene, the foreland basin was separated into several smaller structural and sedimentary basins by rising Laramide uplifts. Subsidence of the Piceance Basin began in the Late Cretaceous (late Campanian) and ended during the middle Eocene (Johnson and Finn, 1986; Johnson, 1990; Johnson and Roberts, 2003). In contrast, subsidence of the Uinta Basin did not begin until the Paleocene and continued into the late Eocene and possibly early Oligocene (Johnson and Finn, 1986).

During the Late Cretaceous, the area that is now the Piceance and Uinta Basin province was located near the western shoreline of the Western Interior Cretaceous seaway (Hettinger and Kirschbaum, 2002). Sediments eroded from the Sevier orogenic belt in eastern Nevada and western Utah formed a broad piedmont of coalesced alluvial fans that graded eastward into alluvial plain, coastal plain, deltaic, and
marine settings that compose the strata of the Mesaverde Group (Cole, 2008; Stancel et al., 2008). This study follows the stratigraphic terminology of Hettinger and Kirschbaum (2002, 2003) for the Mesaverde Group in the southern part of the Piceance and Uinta Basins (Figure 2). This study focuses on the upper part of the Mesaverde Group in both the Piceance and Uinta Basins, which includes the similarly deposited Williams Fork Formation and Farrer and Tuscher Formations, respectively (Lawton, 1986; Hettinger and Kirschbaum, 2002). The Williams Fork Formation is 3600–5155 ft (1097–1571 m) thick along the Grand Hogback and thins westward to approximately 1200 ft (366 m) at the Colorado–Utah state line (Hettinger and Kirschbaum, 2002). In the southwestern Piceance Basin, the Williams Fork Formation is informally subdivided into lower (sandstone-poor), and middle–upper (sandstone-rich) intervals based on lithofacies, architectural elements, and net-to-gross ratio (Cole and Cumella, 2005; Pranter and Sommer, 2011; Keeton, 2012). As observed in Coal Canyon, the lower Williams Fork Formation, which also contains the Cameo–Wheeler coal zone at its base, is a relatively low net-to-gross ratio system (15% average net-to-gross ratio) that largely consists of mudrock and isolated channel form sandstone bodies (channel bars). This lower interval is interpreted to have been deposited by anastomosing to meandering streams in a mostly coastal plain setting to marginal marine setting (Lorenz and Rutledge, 1987; Johnson, 1989; Hemborg, 2000; Patterson et al., 2003; Cole and Cumella, 2005; Pranter et al., 2007). The middle and upper Williams Fork Formation are distinguished by a relatively high net-to-gross ratio system (50%–80% average net-to-gross ratio) containing abundant amalgamated sheet-like channel-form sandstone bodies and associated mudrocks that are interpreted to have been deposited in a low-sinuosity braided alluvial plain environment (Patterson et al., 2003; Cole and Cumella, 2005; German, 2006).

In the Uinta Basin, the Farrer and Tuscher Formations successively overlie the Neslen Formation, a prograding delta complex that includes tidal and coastal plain deposits, and are overlain unconformably by the Paleogene Wasatch Formation (Stancel et al., 2008). The Farrer Formation extends west from the Utah–Colorado border to Soldier Canyon where it grades into the laterally equivalent Price River Formation. It has been measured at 950 ft (290 m) thick at Tusher Canyon in the southeast part of the basin and thins westward to 131 ft (40 m) at its western limit. The Farrer Formation consists of fining-upward single-story sandstone bodies, multistory sandstone bodies, and thick siltstone sequences that are interpreted to have been deposited by a moderate-sinuosity meandering fluvial system in an upper coastal plain environment (Lawton, 1986; Hettinger and Kirschbaum, 2002). Net-to-gross ratio begins to increase in the upper part of the Farrer Formation, and the gradational contact with the overlying Tuscher Formation was placed by Lawton (1983, 1986) to be where the sandstone content exceeded 50%. The Tuscher Formation, which extends westward from the Utah–Colorado border to near Green River, Utah, ranges from 919 ft (280 m) in thickness at Tuscher Canyon to 358 ft (109 m) at its western limit (Lawton, 1986; White et al., 2008). The sandstone dominant Tuscher Formation is characterized by thick amalgamated sheetlike sand bodies with thin siltstone intervals, as well as interspersed thick laterally discontinuous sandstone bodies that feature lateral-accretion surfaces. The Tuscher Formation is interpreted to have been deposited by northeast-flowing meandering and braided fluvial systems (Lawton, 1986).

**METHODOLOGY AND DATA SET**

**Study Area**

Data used in the training and testing of electrofacies models come from 1626 samples that are associated with known architectural element classifications as determined from the detailed sedimentologic analysis of cores (N = 9, total footage 1692 ft [516 m]). Cores were selected for the study on the basis of (1) accessibility, (2) geographic distribution, (3) stratigraphic distribution, and (4) quality (length of core, continuity of cored intervals, and few rubblized zones) (Figures 1, 3). Each sample is also associated with as many as four available measured properties, which consist of the wireline log curves (1) GR, (2) RHOB, (3) deep resistivity (ILD), and (4) NPHI. These measured properties were selected based on their mutual presence for all of the study cores. To evaluate the effective prediction of electrofacies classes in noncored wells,
Figure 2. Schematic cross section showing the stratigraphic relationships and nomenclature of the Piceance and Uinta Basins. The study interval focuses on the upper part of the Mesaverde Group consisting of the Farrer and Tuscher Formations (Fm.) in the Uinta Basin and the Williams Fork Fm. in the Piceance Basin (modified from Hettinger and Kirschbaum, 2002, courtesy of US Geological Survey). Ls. = limestone; Maastricht. = Maastrichtian; Mbr. = member; Paleo. = Paleocene; Sh. = shale; Ss. = sandstone.
the study cores were divided into two subsets: a training set (5 cores, 440 samples) and a testing set (4 cores, 1186 samples).

To demonstrate the applicability of the methods investigated in this study, additional noncored wells ($N = 216$) were selected throughout the study area in which the most successfully trained classifier would be employed to exhibit the batch prediction of electrofacies classes (Figure 1). These wells were visually inspected and selected on the basis of (1) geographic coverage, (2) stratigraphic coverage, (3) robustness of wireline log curve assemblage, (4) data quality (few obvious data spikes, few borehole breakouts, and modern wireline logs), and (5) nondeviated well paths through the interval of interest.

Figure 3. Locations of cored intervals (black bars) in cored wells. Well locations are shown on Figure 1. Wells have equal spacing; no horizontal scale is used. Thickness values are in feet. Fm = formation; GR = gamma ray log (API units); ILD = deep induction log (ohm m); RHOB = bulk density log (g/cm$^3$); Ss. = sandstone.
Data Editing and Normalization

Prior to the creation and testing of the training models, several preprocessing subtasks were performed on the wireline data to help ensure reliable electrofacies class assignments: (1) core-to-log depth shift corrections were made to all well log curves; (2) well log curves were visually inspected to remove obvious data errors or spikes; (3) where necessary, logs were resampled to a common increment of 0.5 ft (0.15 m); (4) and GR curves were normalized using two-point histogram shifting.

Classifying Architectural Elements

Detailed sedimentologic descriptions (lithology, grain size, texture, sedimentary structures, and contacts) of the study cores were used to determine architectural elements to investigate their potential to be grouped into distinct electrofacies classes. To capture the detail necessary for architectural element description, a relatively fine-scale description of lithofacies is required. Eleven lithofacies were described in the study cores on a 0.39-in. (1-cm) basis and include (1) highly fissile mudstone; (2) laminated to mottled mudstone; (3) carbonaceous mudstone; (4) convoluted sandy siltstone; (5) wavy laminated sandy siltstone; (6) argillaceous siltstone; (7) convoluted silty sandstone; (8) ripple laminated sandstone; (9) low- to high-angle cross-bedded sandstone; (10) convoluted sandstone; and (11) structureless, cryptically bioturbated sandstone. These lithofacies are similar to those described by past workers of the fluvial deposits of the Mesaverde Group (Cole and Cumella, 2005; Harper, 2011; Keeton, 2012; Sloan, 2012) and are similar to lithofacies that are universally recognized in fluvial deposits because of the common physical processes that control deposition of clastic fluvial lithofacies (Miall, 1978, 2010). Architectural elements were classified through recognition of distinctive assemblages of these lithofacies in addition to the nature of lower and upper bounding surfaces, internal geometry, and scale (thickness) (Miall, 1985). Based on these characteristics as observed in the study cores and by comparing them to observations made in outcrop in past studies of the fluvial deposits of the Mesaverde Group, the following architectural elements were interpreted:

1. floodplain
2. crevasse splays
3. single-story channel bodies
4. multistory channel bodies
5. amalgamated channel bodies

(Figure 4, Table 1).

Floodplain deposits represent the dominantly fine-grained lithologies surrounding ancient fluvial channels and meander belts and are dominated by the highly fissile mudstone, mottled mudstone, and carbonaceous mudstone lithofacies (Bridge, 2006). Abundant carbonaceous root traces sometimes associated with iron-oxide motting were commonly observed in floodplain deposits throughout the study cores suggesting humid climatic conditions with seasonal flash-flooding and a fluctuating water table (Retallack, 2001; Flaig et al., 2011). Pedogenic features such as small calcite nodules do occur within the floodplain facies assemblage but are very rare, suggesting little time for soil development in a rapidly aggrading floodplain setting (Smith and Rogers, 1999). Floodplain deposits commonly exhibit relatively high GR values, low effective porosity, and bulk density values that are less than sandstone-rich architectural elements.

Crevasses splays develop during high-runoff events when the channel bank is breached and sediment spreads out to be deposited onto the floodplain area (Bridge and Tye, 2000; Miall, 2010). Crevasses splays tend to feature a coarsening-upward grain size trend that typically grades from basal mud-rich floodplain lithofacies into wavy to convolute sandy siltstones and up into very fine- to fine-grained ripple laminated sandstones. These deposits are commonly capped by a sharp contact with overlying floodplain lithofacies and commonly feature a high amount of bioturbation in the form of rooting and probable insect burrowing. Crevasses splays of the Mesaverde Group have been observed in outcrop at Coal Canyon, Colorado, to have an average thickness of 3 ft (0.85 m) (Cole and Cumella, 2005). Crevasses splays commonly exhibit relatively higher GR values as compared with point bar deposits. The GR values can decrease upward and reflect the progradational nature of the deposit. They commonly have lower effective porosity and bulk density values that are less than sandstone-rich architectural elements.

Single-story channel bodies typically represent isolated point bar deposits, which are common in high-
Figure 4. Classification of sandstone body types of the Williams Fork Formation. (A) Crevasse splay, (B) single-story channel body, (C) multistory channel body, and (D) amalgamated channel complex examples are shown with a schematic map view of the depositional setting and cross sectional view of the body. Classification is modified from Cole and Cumella (2005). Figure is modified from Pranter et al. (2009), with permission of AAPG.
sinuosity rivers (Ellison, 2004; Cole and Cumella, 2005; Pranter et al., 2007, 2009; Miall, 2010; Pranter and Sommer 2011). These deposits commonly fine upward and consist of fine- to medium-grained cross-bedded to ripple-laminated sandstone lithofacies with mudchip lags at the base sourced by nearby cutbank erosion. As measured in Mesaverde Group outcrops at Coal Canyon, Colorado, single-story channel bodies average 9 ft (3 m) in thickness (Cole and Cumella, 2005). Point bar deposits commonly exhibit relatively lower gamma-ray values as compared with other architectural element. The GR values can decrease upward and reflect the upward-fining character of the deposit unless several deposits are amalgamated. Single-story channel bodies commonly have relatively higher effective porosity and bulk density values that are slightly greater than crevasse splays and floodplain deposits.

Multistory channel bodies consist of fine- to medium-grained cross-bedded to ripple-laminated sandstone lithofacies, both of which have the potential to be convoluted or bioturbated. These deposits are characterized by stacked individual channels that are made recognizable by the presence of typically approximately two to five internal scours, which may have a mudchip lag present at their base. This complex internal architecture suggests that these channel bodies, which are observed in outcrop at Coal Canyon, Colorado, to average 14 ft (4 m) in thickness, were deposited by dynamic fluvial channels in a robust sinuous channel system with significant meander belts (Cole and Cumella, 2005). Amalgamated channel bodies are composed of stacked assemblages of single-story and multistory channel bodies. In Mesaverde Group outcrops at Plateau Creek Canyon, Colorado, they have been observed to be thicker

<table>
<thead>
<tr>
<th>Architectural Element Class</th>
<th>Principle Facies Assemblage</th>
<th>Description</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floodplain</td>
<td>MF, ML, MC, STAM</td>
<td>Mud-dominated facies assemblage that typically features high degrees of bioturbation in the form of carbonaceous root traces, which are commonly iron-stained.</td>
<td>Floodplain</td>
</tr>
<tr>
<td>Crevasse splay</td>
<td>STSC, STWR, SSTC, SR</td>
<td>Typically, a coarsening-upward facies sequence with a gradational basal contact and a sharp overlying contact. Bioturbation in the form of rooting and insect burrowing is common.</td>
<td>Crevasse splay</td>
</tr>
<tr>
<td>Single-story channel body</td>
<td>SL, SC, SS, SR</td>
<td>Singular, fine- to medium-grained sand-dominated fining-upward sequence with sharp basal contact commonly featuring mud clasts.</td>
<td>Isolated point bar deposit in high-sinuosity fluvial system</td>
</tr>
<tr>
<td>Multistory channel body</td>
<td>SL, SC, SS, SR</td>
<td>Consists of multiple stacked individual channel bodies made recognizable by the presence of between two and five internal scour surfaces.</td>
<td>Stacked individual channels in a robust sinuous channel fluvial system</td>
</tr>
<tr>
<td>Amalgamated channel body</td>
<td>SL, SC, SS, SR</td>
<td>A complex consisting of stacked multistory and single-story channel bodies typically much larger in scale than a single multistory channel body.</td>
<td>Fluvial channel deposits in a low- to medium-sinuosity braided fluvial system</td>
</tr>
</tbody>
</table>

Abbreviations: MC = carbonaceous mudstone; MF = highly fissile mudstone; ML = laminated to mottled mudstone; SC = convoluted sandstone; SL = low- to high-angle cross-bedded sandstone; SR = ripple laminated sandstone; SS = structureless, cryptically bioturbated sandstone; SSTC = convoluted silty sandstone; STAM = argillaceous siltstone; STSC = convoluted sandy siltstone; STWR = wavy laminated sandy siltstone.
(average thickness of 26 ft [8 m]) and more laterally extensive (average width of 870 ft [265 m]) than multistory channel bodies and are interpreted to have been deposited by a low- to medium-sinuosity braided river system in an alluvial plain setting (Lorenz and Nandon, 2002; Patterson et al., 2003; German, 2006).

**Measuring Success of Classification Methods**

A means of judging the trained classifiers’ ability to correctly predict electrofacies classes in test wells is needed to evaluate and compare classifier performances. A commonly used visualization tool, typically used in evaluating supervised learning procedures in the field of artificial intelligence, known as a confusion matrix was selected for this purpose (Ting, 2011). A confusion matrix displays information about the actual and predicted classifications present in a classification system, where each column represents instances of the predicted class and each row represents instances of the actual class (Kohavi and Provost, 1998; Ting, 2011). Data organized in this manner not only display correctly predicted classes but also allow for characterization of erroneous interclass predictions (Foody, 2002).

To quantitatively compare performances between variably trained classifier outcomes, four metrics of success were derived from the content of the confusion matrix: overall accuracy, user’s accuracy, predicted volume, and average deviation. Overall accuracy is the simplest and one of the most common accuracy measurements used in confusion matrix analysis (Congalton, 1991). It is calculated by dividing the number of correctly predicted classes (the sum of the major diagonal in the confusion matrix) by the total number of predicted classes. To measure the accuracy of individual class predictions, a similar calculation to overall accuracy is derived from the confusion matrix whereby the number of correctly predicted instances of a particular class is divided by the total number of actual samples that exist for that class. This success metric is known as user’s accuracy and is commonly employed in conjunction with overall accuracy when an emphasis is placed on the accuracy of individual class predictions as is the case in this study (Janssen and van der Wel, 1994; Foody, 2002). In addition to measures of accuracy, it is also desirable that the number of predicted samples of a particular class is similar to the number of actual samples for that class so that the model accurately represents the actual volumetric distribution of facies (Dubois et al., 2007). To evaluate the volumetric distributions of individual classes, the total number of predicted samples for a particular class is divided by the number of actual samples for that class. This ratio is multiplied by 100 to give a measure in terms of percentage of the correctness of the predicted volumes of each individual class from the ideal percentage (100%). This measure of central tendency is known as the average deviation.

**The k-Nearest Neighbor Approach**

In the k-NN algorithm, the training phase consists simply of assigning and storing class labels to training samples, which are vectors in n-dimensional space. In the classification phase, an unclassified sample (a test sample or query point) is plotted among the training data in n-dimensional space and is compared with a user-defined constant number (k) of the most similar training samples (nearest neighbors). The test sample is then classified according to the most commonly occurring class out of this k number of nearest training samples (majority rules) (Cover and Hart, 1967; Dubois et al., 2007) (Figure 5). This algorithm is similar to, but not to be confused with, the popular k-means clustering algorithm.

Training samples (N = 440) with known architectural element classifications were selected at random throughout the four training cores where data quality was deemed satisfactory (high quality of associated wireline data, reliable core-to-log depth shift correction, and confidently chosen architectural element classifications). The n-dimensional space that the training samples are plotted in corresponds to the number of different well log curves that are used, with each sampling point being associated with the well log values at its respective depth. To investigate the effectiveness of the well log curves chosen for this study (GR, RHOB, ILD, and NPHI) in distinguishing between architectural element classes, the performance of the classifier as trained by seven different well log curve assemblages was evaluated. These seven well log
Curve assemblages consist of (1) GR and RHOB; (2) GR and ILD; (3) GR and NPHI; (4) GR, RHOB, and ILD; (5) GR, RHOB, and NPHI; (6) GR, NPHI, and ILD; and (7) GR, RHOB, ILD, and NPHI. The GR log curve was left static in these assemblages because it is present in all of the selected noncored study wells.

Determining the number of nearest neighbors \( k \) to examine when classifying the test sample can be a delicate choice. If the value of \( k \) is too small, there is the potential for outlying sampling points to have a greater influence on test sample classification. If the value of \( k \) is too large, there is the tendency for classes that are associated with larger sampling populations to have an overwhelming influence on test sample classification (Drummond et al., 2010). Because there is not a widely accepted formula for determining the optimum \( k \) value and attempts at such are complicated, the example of Drummond et al. (2010) was followed where a series of 5 different \( k \) values (5, 10, 15, 20, and 25) were tested (Hall et al., 2008). These \( k \) values were used in conjunction with 7 different log curve assemblages to create 35 uniquely trained classifier models. After initial evaluation of the trained classifiers’ ability to predict the five original architectural element classes as described in core, a simplified four-class architectural element realization was created and tested in the manner described above.

**Probabilistic Clustering Procedure**

In the PCP, an initial model is created wherein core-defined architectural element classes and their associated well log curve values are stored as sampling points. For consistency, the same core depth sampling points as well as the same well log curve assemblages used in training the \( k \)-NN classifier were used to train the PCP classifier. In the imposed model, the frequency distribution of each selected well log curve is segregated into a user-defined number of pseudo-Gaussian distributions that represent each desired electrofacies class \((N = 4 \text{ and } 5 \text{ in this study})\). These pseudo-Gaussian distributions are plotted as clusters or modes in \( n \)-dimensional space, with \( n \) being dependent on the number of well log curves used (Vallejo, 2010; Eslinger and Everett, 2012). During the clustering process a probability density function is employed to calculate the likelihood that each sample belongs to a particular electrofacies class. The sampling points obtained from core description that are used in the calibration of the model are initially assigned a probability of 1.0. Data from the testing wells were withheld from the calibration process so that they would not influence the initial model. To incorporate the data from the test cores so that electrofacies class predictions can be made, at least one adjusting calculation (one iteration) must be made in which a new probability density function is computed. This iteration not only assigns electrofacies class probabilities in the test cores but may also change the class assignment probabilities of the training samples. Generally, when greater numbers of iterations are computed, there exists a higher chance that the training samples will be reassigned to a new electrofacies class (Vallejo, 2010; Eslinger and Boyle, 2011). For this reason, all of the models tested using the PCP in this study were only allowed the minimum requirement of one iteration. The final classification given to samples in the testing wells is determined by the electrofacies class that is calculated to have the highest probability.

**Thickness Criteria Approach**

Past studies of the Mesaverde Group in the Piceance Basin have established relationships between the fluvial...
channel body and crevasse splay architectural elements and their respective thicknesses (e.g., Cole and Cumella, 2005; Pranter et al., 2009).

These relationships were incorporated in the refinement of classification results by applying a method involving an indicator flag, which pairs the results of the electrofacies classifiers with thickness criteria. For the thickness criteria approach, the thickness of a user-defined well log value is counted from top to bottom, whereas unspecified values or gaps can be overlooked if they fall below a predefined thickness limit. If the well log value being counted passes a gross thickness requirement (thickness including gaps that fall below the predefined thickness limit) and a net thickness requirement (thickness of user-defined value excluding gaps), then an indicator flag is created denoting a newly assigned class (Figure 6). This method was applied to the k-NN classification results for the four-class architectural element realization to help distinguish between the crevasse splay, single-story channel body, and multistory channel body classes. Thickness criteria are based on average thickness values of architectural elements as observed in outcrop at Coal Canyon, Colorado, by Cole and Cumella (2005). The workflow for the thickness criteria approach is as follows. (1) Test wells in which electrofacies have been predicted by the k-NN classifier are used (as log curves in .las file format) with each predicted electrofacies corresponding to a representative value (1 = floodplain, 2 = crevasse splay, 3 = single-story channel body, and 4 = multistory channel body). (2) A multistory channel flag is created using a gross thickness requirement of 12 ft (4 m), which is set at roughly 2 ft (0.61 m) less than the average thickness of a multistory channel body (14 ft [4 m]) such that the thickness requirement is slightly more inclusive than the average thickness value (Figure 6). A net thickness requirement of 9 ft (3 m) was deemed appropriate based on visual inspection of classifier results in well log form, and the maximum acceptable gap was set at 4 ft (1 m) based on previous use of this method (Figure 6). (3) After the multistory channel body flag has been established, a three-class electrofacies log curve is created, which consists of (i) the floodplain electrofacies; (ii) a new class that combines the crevasse splay electrofacies, the single-story channel body electrofacies, and any interval that was previously classified as a multistory channel body but did not pass the thickness requirements necessary to be considered part of the multistory flag; and (iii) the multistory channel body flag just created. (4) A single-story channel body indicator flag is then applied to this new simplified electrofacies log curve using a gross thickness requirement of 7 ft (2 m), which, like was done for the multistory channel body indicator flag, is set at approximately 2 ft (0.6 m) less than the average thickness of a single-story channel body (9 ft [3 m]) such that the thickness requirement is slightly more inclusive than the average thickness value. A net thickness value of 4 ft (1 m) was chosen based on visual inspection of classifier results, and the maximum acceptable thickness gap was kept to 0.5 ft (0.15 m) so as to eliminate any small erroneous classification spikes while not allowing the grouping of closely stacked similar electrofacies. Within this three-class electrofacies log curve, if an interval of values belonging to the second class (ii) does not meet the minimum thickness requirements to be flagged as a single-story channel body, it is classified as a crevasse splay. (5) A final electrofacies log curve is created that merges the two channel flags with the floodplain and crevasse splay classes. A coal flag was also created with the criteria GR < 75 API and RHOB < 2.1 g/cm³ and merged along with these classes into the final electrofacies log curve.

**Figure 6.** Example illustrating the process by which a well log indicator flag for a multistory channel body is created.
RESULTS

Through an approach that combined selected classifier results with the thickness criteria approach, an overall accuracy of 83.6% was achieved. The individual architectural elements of the simplified four-class architectural element realization were predicted with user’s accuracies of 0.891, 0.376, 0.735, and 0.985 for the floodplain, crevasse splay, single-story channel body, and multistory channel body classes, respectively. Such a methodology could provide an alternative approach to the classification of fluvial architectural elements in commonly available well logs (Figure 7).

Prediction Performance of the $k$-Nearest Neighbor Method

Evaluation of the variably trained (varying log curve assemblages and $k$ values) $k$-NN classifiers’ ability to correctly predict the occurrence of the five architectural elements as they exist in the testing wells rendered a best case overall accuracy of 63.8% with the floodplain, crevasse splay, single-story channel body, multistory channel body, and amalgamated channel body yielding individual user’s accuracies of 0.897, 0.624, 0.026, 0.228, and 0.904, respectively (Figure 8). Although the $k$-NN classifier fared quite well in its ability to predict the floodplain, crevasse splay, and amalgamated channel body architectural elements, the single-story channel body and multistory channel body architectural elements were rarely predicted correctly. An examination of interclass confusion in all of the $k$-NN tests revealed a high degree of confusion between the single-story channel body and crevasse splay architectural elements with 132 out of the 189 total single-story channel body samples being incorrectly classified as the crevasse splay class. A high degree of confusion between the multistory channel body and amalgamated channel body architectural element was also noted with 92 out of 158 multistory channel body samples being incorrectly classified as the amalgamated channel body class. This confusion is also represented by the high predicted volumes of the crevasse splay and amalgamated channel body classes and the low predicted volumes of the single-story channel body and multistory channel body classes, which contributed to an average deviation value of 62.6 (Figure 8). A new simplified four-class architectural element realization was developed in which the highly confused and geologically similar multistory channel body and amalgamated channel body architectural elements were combined into the same multistory channel body class. Evaluation of the $k$-NN classifier’s ability to predict the simplified four-class realization resulted in an improved best case overall accuracy of 74.5% with the floodplain, crevasse splay, single-story channel body, and multistory channel body yielding individual user’s accuracies of 0.926, 0.539, 0.069, and 0.932, respectively (Figure 9). Simplification of the architectural element classes also

Figure 7. The well log curve of the far right column features the resulting architectural elements as predicted in a section of test well, Kerr–McGee 21 Natural Buttes, by the coupling of the well log indicator flag approach with the results of the $k$-nearest neighbor classifier as trained by the gamma ray (GR) and bulk density (RHOB) well log curves. This is compared with a well log curve to its left in which the actual core-described architectural elements are featured.
led to an improvement in interclass confusion marked by a more desirable predicted volume for the newly aggregated multistory class and a consequently lower average deviation (45.9) compared with the previously tested five-class architectural element realization. However, the confusion between the single-story channel body and crevasse splay architectural element classes still remained high with 111 out 189 total single-story channel body samples being incorrectly classified as the crevasse splay class (Figure 9). Observations of the success of the varying well log curve assemblage and $k$ values in distinguishing the original five architectural element classes were ignored in favor of focusing on the significance of these variables in distinguishing between the classes of the more geologically realistic four-class architectural element realization.

The range in overall accuracy values produced by the seven well log curve assemblages tested was not large, and no one well log curve assemblage predicted an individual architectural element class markedly better than the others. However, the well log curve assemblages of GR and RHOB and GR, RHOB, and ILD are associated with the identical highest overall accuracies achieved. The addition of the ILD well log curve seemed to have had little to no effect on the outcome of the predictions, indicating that the GR and RHOB curves are primarily responsible for the success of the predictions (Figure 10). Testing of the five different $k$ values (5, 10, 15, 20, and 25) with the various well log curve assemblages revealed a general increase in overall accuracy up to the $k$ value of 20 followed by a decline in overall accuracy at the $k$ value of 25 in the majority

---

**Figure 8.** (A) User’s accuracies for the five-class architectural-element realization produced by the $k$-nearest neighbor ($k$-NN) classifier trained by the log-curve assemblage gamma ray (GR), bulk density (RHOB), neutron porosity (NPHI), and $k$ value of 20—one of two variably trained $k$-NN classifiers to produce the highest overall accuracy achieved (63.8%) for the five-class architectural element realization. (B) The confusion matrix associated with the $k$-NN classifier that was trained using the well log curves GR, RHOB, true resistivity (RT), and NPHI and a $k$ value of 20, which produced an overall accuracy of 63.8% achieved when attempting to predict the five architectural element classes. Highlighted cells represent the number of correctly predicted samples out of the actual class sum, whereas the other cells within the row represent the misclassified architectural elements. AM = amalgamated channel body; CS = crevasse splay; FP = floodplain; MS = multistory channel body; SS = single-story channel body.

---

**Table:** Example Confusion Matrix for 5-class Architectural-element Realization by k-NN

<table>
<thead>
<tr>
<th>Actual Class</th>
<th>Predicted Class</th>
<th>Actual Class Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FP</td>
<td>CS</td>
</tr>
<tr>
<td>FP</td>
<td>411</td>
<td>46</td>
</tr>
<tr>
<td>CS</td>
<td>32</td>
<td>88</td>
</tr>
<tr>
<td>SS</td>
<td>1</td>
<td>132</td>
</tr>
<tr>
<td>MS</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>AM</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Column Total</td>
<td>444</td>
<td>288</td>
</tr>
</tbody>
</table>

Predicted Volume | 97% | 204% | 9%  | 41% | 155% |
of cases (Figure 11). This trend in overall accuracy is mirrored by the user’s accuracies of the individual architectural elements with the exception of the single-story channel body, which commonly decreases in user’s accuracy with successively larger \( k \) values. The classifiers trained by the combination of these optimal variables of \( k = 20 \) and the well log curve assemblage of GR and RHOB are responsible for producing the best case results for the four-class architectural element realizations that were previously mentioned.

**Prediction Performance of the Probabilistic Clustering Procedure Classifier**

Like the \( k \)-NN method, when evaluated for its ability to distinguish the original five architectural elements, the PCP proved capable of predicting the floodplain, crevasse splay, and amalgamated channel body architectural elements with moderate to high user’s accuracies, which contributed to a best case overall accuracy of 62%. However, the single-story channel body and multistory channel body were rarely predicted correctly by the PCP (Figure 12). Analysis of interclass confusion revealed that the single-story channel body was commonly incorrectly predicted as either the crevasse splay or amalgamated channel body architectural elements. A high degree of confusion between the multistory channel body and amalgamated channel body architectural elements was also observed with 93 out of 158 multistory channel body samples being incorrectly classified as the amalgamated channel body class. This confusion...
is also represented by the high predicted volumes of the crevasse splay and amalgamated channel body classes and the low predicted volumes of the single-story and multistory channel body classes, which contributed to an average deviation value of 52.6 (Figure 12). As was done during the testing of the $k$-NN method, the PCP was tested for its ability to predict classes in a simplified four-class architectural element realization, which combines the highly confused and geologically similar multistory channel body and amalgamated channel body classes. Once again, the simplified four-class architectural element realization lent itself to the alleviation of some degree of interclass confusion denoted by a lower average deviation value (42.7) attributed to more ideal predicted volumes of the newly aggregated multistory class (Figure 13). Though success metrics for the multistory class improved, the single-story channel body is still almost never predicted correctly (Figure 13). Of the seven well log curve assemblages tested, none produced dramatically better overall accuracies than the others; however, although the difference is small, the well log curve assemblages GR and RHOB and GR, RHOB, and ILD are associated with the highest overall accuracies achieved by the PCP classifier of 72.8% (Figure 10). Again, the ILD well log curve seemed to have little to no effect on the accuracy of the predictions suggesting the GR and RHOB well log curves are primarily responsible for the success of the predictions. With the PCP classifier, none of the seven well log curve assemblages showed an advantage over the others in their ability to predict the individual architectural element classes with the exception that those assemblages that included the RHOB curve were more effective at predicting the crevasse splay architectural element class than those that did not include it.

**Refinement of Classifier Results with the Thickness Criteria Approach**

Though both classifiers were shown to be capable of predicting the floodplain, crevasse splay, and multistory channel body classes with satisfying accuracy in the simplified four-class architectural element realization, attempts by both classifiers failed to accurately predict the single-story channel body class, which was always highly confused for the crevasse splay and/or multistory channel body classes. After applying the thickness criteria approach to the classifier result having the best overall accuracy ($k$-NN classifier trained by well log curve assemblage GR and RHOB and $k = 20$), the previously troublesome single-story channel body class was now predicted with a 0.735 user’s accuracy, bringing the overall accuracy to
a new high of 83.6%. The improvement in accuracy metrics is owed to a reduction in interclass confusion made possible by the incorporation of thickness criteria. The single-story channel body class was now much less commonly misclassified as the crevasse splay class, a change that is reflected in more ideal predicted volumes and average deviation value (Table 2). The user’s accuracies of the other architectural element classes remained relatively unchanged compared with their success in the k-NN method, with the exception of the crevasse splay, which decreased slightly. A depiction of the well log curves generated by the thickness criteria approach as compared with well log curves representing architectural elements as described in the test cores is shown in Figure 14.

**DISCUSSION**

**Combining Alike Architectural Elements**

The resulting architectural element predictions of the two classification methods displayed many similarities in terms of the accuracies that were achieved as well as the interclass confusion that was revealed upon examination of the confusion matrices. When tested for their ability to distinguish the five architectural elements that were originally described in core, both the k-NN and PCP classifiers showed a high degree of confusion between the multistory channel body and amalgamated channel body classes. In this study, the use of the term amalgamated channel body was influenced by the observations of this class by Cole and Cumella (2005) at Main Canyon and Plateau Creek Canyon, Colorado, and further detailed study by German (2006) where it was described as the stacked multistory channel bodies that were thicker (average thickness of 26 ft [8 m]) and more laterally extensive (average width of 870 ft [1265 m]) than a typical multistory channel body. Beyond scale, which can be difficult to accurately assess in core, the geologic similarity of these two architectural element classes made their distinction in core a subjective process. The subjectivity of this distinction is likely reflected in the high degree of confusion observed between the two classes. Combining the similar multistory channel body and amalgamated channel body architectural elements into a single class to create a simplified four-class architectural element realization seemed a geologically realistic choice given that both of these architectural elements can potentially be deposited by low- to medium-sinuosity fluvial systems (White et al., 2008). The single-story channel body class was found to be highly confused with the crevasse splay class by both classifiers, and combining the two classes would have likely also led to improved measures of interclass confusion. However, these architectural elements were kept as discrete classes because of their unique depositional and reservoir characteristics, which would carry significance if the methods used in this study were to be applied in future mapping and modeling endeavors.

**Comparing the Classifiers as Trained by the Different Well Log Curve Assemblages**

Evaluation of the effectiveness of the two classifiers as trained by the different well log curve assemblages was based upon their ability to predict the four-class architectural element realization. It was reasoned that measures of accuracy would be more meaningful for this more geologically realistic four-class realization as
compared with the five-class architectural element realization. In both the \( k \)-NN and PCP classifiers, the seven different well log curve assemblages tested resulted in a small range of overall accuracies achieved by both classifiers with the log curve combination of GR and RHOB being primarily responsible for modestly higher overall accuracies achieved (Figure 10). To avoid the pitfall of making inferences of the effectiveness of the well log curve assemblages based only on the overall accuracy metric, which can potentially be misleading if a large testing group is very well predicted (e.g., the floodplain or multi-story channel body testing groups), user’s accuracy of the individual architectural elements must also be inspected. Any differences in the user’s accuracies of the individual architectural elements produced by the different well log curve assemblages in the \( k \)-NN classifier are negligible; they are just slightly more accurate when trained by the well log curve assemblage GR and RHOB as suggested by the overall accuracy metric. The agreement between the overall accuracy and user’s accuracy metrics produced by the different well log curve assemblages suggests that the log curve assemblage does not make an immense difference but is slightly more accurate with the combination of GR and RHOB. This is a beneficial revelation for possible future application of this method in the fluvial deposits of the Mesaverde Group.
where well log curve suites can vary substantially. When user’s accuracy was evaluated in the outcomes of the PCP classifier, it was found that the well log curve assemblages that included the RHOB well log curve were roughly 20% more effective at predicting the crevasse splay class than those that did not contain RHOB. The approximately 20% swing observed in the user’s accuracy of the crevasse splay class while the overall accuracy concurrently responds to a much lesser degree highlights the potential for disparity between the two metrics and the need for the simultaneous evaluation of both. A study in the nearby Mamm Creek field, Piceance Basin, Colorado, also emphasized GR and RHOB as the most important well log curve assemblage in a method devised to predict diagenetic facies in a single well experiment in the upper part of the Mesaverde Group (Ozkan et al., 2011). The authors speculated that aside from diagenetic facies, the method may also be useful in predicting the crevasse splay architectural element through its recognition of fine-grained and sometimes clay matrix-rich siltstones and sandstones (Ozkan et al., 2011).

As a whole, the top overall accuracy for each well log curve assemblage for the $k$-NN classifier was, in most cases, slightly higher than the overall accuracy produced by the same well log curve assemblage in the PCP classifier (Figure 10); however, the results are similar. The $k$-NN classifier also did not show the same RHOB well log curve–dependent contrast in its ability to predict the crevasse splay class as was observed in the PCP

![Figure 13. (A) User’s accuracies for the four-class architectural element realization produced by probabilistic clustering procedure (PCP) classifier trained by the log curve assemblage gamma ray (GR) and bulk density (RHOB)—one of two log curve assemblages to produce the highest overall accuracy achieved (72.8%) by the PCP classifier for the four-class architectural element realization. (B) The confusion matrix associated with the PCP classifier that was trained using the well log curve assemblage GR and RHOB, which produced an overall accuracy of 72.8% achieved when attempting to predict the four architectural element classes. Highlighted cells represent the number of correctly predicted samples out of the actual class sum, whereas the other cells within the row represent the misclassified architectural elements. CS = crevasse splay; FP = floodplain; MS = multistory channel body; SS = single-story channel body.](image)

<table>
<thead>
<tr>
<th>Predicted Class</th>
<th>Actual Class Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Class</td>
<td>FP</td>
</tr>
<tr>
<td>FP</td>
<td>413</td>
</tr>
<tr>
<td>CS</td>
<td>36</td>
</tr>
<tr>
<td>SS</td>
<td>11</td>
</tr>
<tr>
<td>MS</td>
<td>0</td>
</tr>
<tr>
<td>Column Total</td>
<td>460</td>
</tr>
<tr>
<td>Predicted Volume</td>
<td>100%</td>
</tr>
</tbody>
</table>

![User’s Accuracy of 4-class Architectural-element Realization by PCP](image)
classifier. Similar to the findings of this study, the comparative strength of the \( k \)-NN classifier to other commonly used classifiers was also reported by Dubois et al. (2007).

**The Troublesome Single-Story Channel Body Architectural Element Class**

The main source of interclass confusion in the four-class architectural element realization seen in both classifiers was caused by the consistent misclassification of the single-story channel body architectural element class. In the \( k \)-NN classifier it was observed that the troublesome class was better (though still poorly) predicted at low \( k \) values, and its accuracy became progressively poorer with each successive increase in \( k \). A commonly recognized fault in the \( k \)-NN algorithm is that if the \( k \) value used is too large, the classes with larger training populations have the potential to overwhelm classes with smaller training populations (Drummond et al., 2010). To inspect if the relatively smaller training population of the single-story channel body class was having a negative effect on the success of its prediction, a test was conducted wherein the training populations of the different classes were equalized (\( N = 141 \) for each class), thus taking away the potential for class size influence. This resulted in a very slight increase in user’s accuracy values (<6% improvement over the previous top user’s accuracy) for the single-story channel body. This minimal improvement indicates that if there were a training population sample size influence, it was likely insignificant. Equalization of the training populations was also found to be unbenevolent to the success of lithofacies prediction using the \( k \)-NN classifier in a study by Dubois et al. (2007). Examination of 50 random training samples each of the crevasse splay and single-story channel body classes cross-plotted against the useful well log curves of GR and RHOB illustrates the degree of similarity between the two classes and helps explain why the potential existed for such confusion. Cross-plotting of training samples of the single-story channel body class and the multistory channel body class also illustrates the potential for confusion among these classes.

Because diagenetic and mineral composition vary both geographically and stratigraphically in the fluvial deposits of the Mesaverde Group (Keighin and Fouch, 1981; Johnson and Roberts, 2003; Stroker et al., 2013), without conducting detailed petrographic work it is only possible to broadly speculate on geologic drivers for the single-story channel body interclass confusion. It was observed in core that single-story channel bodies could range from upper very fine to medium sand grain size. This range in grain size overlaps to some degree with both the crevasse splay grain sizes (silt to fine sand) and the multistory channel body grain sizes (upper fine to medium upper sand). All three of these architectural elements also contain the ripple-laminated sandstone lithofacies to varying extents. These common characteristics are also recorded in outcrop descriptions of Mesaverde Group fluvial architectural elements by Cole and Cumella (2005), White et al. (2008), Harper (2011), Hlava (2011), and Keeton (2012). Though well logs do not measure grain size directly and are most certainly blind to details that geologists commonly describe in cores like ripple laminated sandstones, it is possible that these overlapping attributes of grain size and lithofacies are also associated with changes in porosity and clay content. Although

<table>
<thead>
<tr>
<th>Actual Class</th>
<th>Predicted Class</th>
<th>Predicted Class Sum</th>
<th>Actual Class Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP</td>
<td>408</td>
<td>29</td>
<td>7</td>
</tr>
<tr>
<td>CS</td>
<td>34</td>
<td>53</td>
<td>0</td>
</tr>
<tr>
<td>SS</td>
<td>9</td>
<td>35</td>
<td>139</td>
</tr>
<tr>
<td>MS</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Column total</td>
<td>451</td>
<td>123</td>
<td>166</td>
</tr>
<tr>
<td>Predicted volume</td>
<td>98%</td>
<td>87%</td>
<td>88%</td>
</tr>
</tbody>
</table>

The confusion matrix resulting from the application of the well log indicator flag approach in refining the results of the \( k \)-nearest neighbor (\( k \)-NN) classifier that was trained using the well log curves gamma ray and bulk density and a \( k \) value of 20. Gray highlighted cells represent the number of correctly predicted samples out of the actual class sum, whereas the other cells within the row represent the mispredicted architectural elements. The incorporation of thickness criteria greatly alleviated interclass confusion involving the misclassification of the SS architectural element class. This improvement is reflected in the more ideal predicted volumes achieved and the lowered average deviation value of 9.6 as compared with the average deviation value of the pre-refined \( k \)-NN classifier of 45.9. This improvement in interclass confusion led to an overall accuracy of 83.6%.

Abbreviations: CS = crevasse splay; FP = floodplain; MS = multistory channel body; SS = single-story channel body.
**Figure 14.** Actual and predicted architectural elements in testing wells. Testing wells with architectural elements as classified from core descriptions are in the left track for each of the four wells. These are compared with the architectural elements as predicted through the combined approach of integrating select classifier results ($k$-nearest neighbor) with outcrop-based architectural element thickness criteria (right track for each of the four wells).
porosity and clay content may be difficult for the geologist to adequately describe in hand sample, they do affect the response of the GR and RHOB curves used in this study. Perhaps shared similarities in porosity and clay content with parts of the crevasse splay and multistory channel body classes could have contributed to the frequent misclassification of the single-story channel body.

**Incorporation of Thickness Criteria and Geologically Constrained Electrofacies Prediction**

The resulting improvement in single-story channel body prediction via the thickness criteria approach makes the methods described in this study a much more viable means of predicting the fluvial architectural elements of the Mesaverde Group in noncored wells. An unfortunate byproduct of the thickness criteria approach was the moderate (16%) decrease in user’s accuracy experienced by the crevasse splay class. This is a consequence of the merging of the crevasse splay and single-story channel body classes prior to running the single-story channel body thickness criteria code described in the methods, which resulted in crevasse splays that directly overlie or underlie a single-story channel body being grouped with the contiguous single-story channel body. This side effect should be kept in mind during future application of this process.

No limit exists to how the thickness criteria of this program can be manipulated to tailor the thickness criteria approach to a particular study. As an example, it may be desirable to attempt the prediction and mapping of the large amalgamated channel bodies, deemed “super stories” by some, which are proposed to represent deposition in incised valleys (White et al., 2008). However, it should be cautioned that there is always the potential for the well in which the architectural elements are being predicted to have penetrated a large sandstone body at its comparatively thinner margin, which may result in its misclassification as an isolated single-story channel body.

To demonstrate the application of the methods described in this paper, the results of the $k$-NN classifier trained by the GR and RHOB well log curves and a $k$ value of 20, which produced the highest overall accuracy of all uniquely trained classifiers (74.5%), were paired with the thickness criteria approach to predict the fluvial architectural elements of the four-class realization in the Mesaverde Group fluvial interval of noncored wells ($N = 216$) throughout the study area. Mapping of the cumulative thickness of the single-story channel body architectural element revealed an area of relatively high cumulative thickness in the northwestern part of the study area (Figure 15). This thickening corresponds to the area in and around Natural Buttes field, Utah, in which highly discontinuous, lenticular sandstone bodies have been recognized as the primary Mesaverde Group reservoirs (Schmoker et al., 1996; Stancel et al., 2008). When cumulative thickness values of the multistory channel body class were mapped, a roughly east-to-west fairway was noted in the southern part of the Piceance Basin, Colorado (Figure 15).

The extent of well control prevents investigation in this southernmost part; however, it is noted in outcrop and well logs that a fairway of large amalgamated channel bodies is present from Plateau Creek Canyon, Colorado, to Parachute field, Colorado (Foster, 2010) (Figure 15). Perhaps the relative increase in cumulative thickness for the multistory channel body class observed in the generated maps is related to this nearby occurrence of large amalgamated channel bodies. The cumulative thicknesses of both the single-story and multistory channel bodies are observed to follow general Mesaverde Group thickness trends where they decrease at the Douglas Creek arch, Colorado, and thicken toward the center of the Piceance Basin, Colorado (Hettinger and Kirschbaum, 2002).

**CONCLUSIONS**

Although previous studies of the fluvial deposits of the upper Mesaverde Group have historically relied on the analysis of core samples, outcrops, or the manual interpretation of well logs to develop architectural element classifications, this study explores an alternative approach through the use of statistical classification methods. Through the combination of selected classifier results with outcrop-based, architectural element thickness criteria, an overall accuracy of 83.6% was achieved for the simplified four-class architectural element realization. The individual architectural elements were predicted with...
user’s accuracies of 0.891, 0.376, 0.735, and 0.985 for the floodplain, crevasse splay, single-story channel body, and multistory channel body classes, respectively. Combining the thickness criteria approach with the selected classifier results signified a vast improvement in the prediction of the single-story channel body architectural element class, which both the k-NN and PCP methods on their own were unable to distinguish as a distinct electrofacies class.

Significant differences in prediction performance were not elicited by the classifiers as trained by the seven different well log curve assemblages with the exception that the PCP classifier predicted the crevasse splay class more competently when the RHOB well log curve was used. Though the disparity was modest, the use of well log curves GR and RHOB was associated with the highest overall accuracies achieved by both the k-NN and PCP classifiers. The common occurrence of these well log curves in the study area bodes well for future use of the studied classification techniques. Unique to the k-NN classifier, the potential for improved prediction performance based on choice of k value was also demonstrated, highlighting the benefits of experimenting with this variable.

When coupled with the outcrop-based thickness criteria approach, the methods described in this study present a feasible approach to the classification of the fluvial architectural elements from well logs for the Mesaverde Group and formations in other basins. This has implications for future work that could allow for the creation of maps and three-dimensional reservoir models to be conducted in a timely fashion with results that are objective and easily reproducible.

REFERENCES CITED


