



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
INGENIERÍA CIVIL – CONSTRUCCIÓN

APLICACIÓN DE LA MINERÍA DE DATOS PARA LA OPTIMIZACIÓN DE LA SELECCIÓN
DE MEZCLAS DE CCR PARA SU USO EN PRESAS DE GRAN ALTURA

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN INGENIERÍA

PRESENTA:
CARLOS GABRIEL PADILLA MEJÍA

TUTOR PRINCIPAL
DRA. SILVIA RAQUEL GARCÍA BENÍTEZ
INSTITUTO DE INGENIERÍA

MÉXICO, D.F. OCTUBRE 2015

JURADO ASIGNADO:

Presidente: M. I. Carlos Javier Mendoza Escobedo

Secretario: Dr. Jesús Hugo Meza Puesto

Vocal: Dra. Silvia Raquel García Benítez

1^{er.} Suplente: M. I. Marco Tulio Mendoza Rosas

2^{d o.} Suplente: M. I. Luis Candelas Ramírez

Lugar o lugares donde se realizó la tesis: México, Distrito Federal

TUTOR DE TESIS:

Dr. Silvia Raquel García Benítez

FIRMA

(Segunda hoja)

Contenido

Contenido	3
Índice de Tablas	5
Índice de Figuras	6
INTRODUCCIÓN	8
1. CAPÍTULO I. CONCRETO COMPACTADO CON RODILLO (CCR) EN PRESAS DE GRAN ALTURA.	10
2. CAPÍTULO II. CARACTERÍSTICAS DEL CCR	19
2.1. Materiales.....	19
2.1.1. Materiales cementantes	20
2.1.2. Agregados	23
2.1.3. Aditivos.....	26
2.2. Selección de la dosificación de la mezcla	27
2.2.1. Permeabilidad.	27
2.2.2. Densidad.	28
2.2.3. Resistencia.....	29
2.2.4. Segregación	30
2.2.5. Trabajabilidad.....	31
2.2.6. Durabilidad.	31
2.2.7. Condiciones constructivas.....	32
2.3. Métodos de dosificación de mezclas.....	32
2.4. Mezclas de prueba en laboratorio	34
3. CAPÍTULO III. APRENDIZAJE MÁQUINA	36
3.1. ¿Qué es el aprendizaje automático o aprendizaje máquina?	36
3.1.1. Aprendizaje de funciones de entrada-salida	38
3.1.2. El aprendizaje y el sesgo.	41
3.2. Árboles de decisión.....	42
3.3. Aprendizaje supervisado de árboles de decisión univariados.....	45
4. CAPÍTULO IV. APLICACIÓN DEL MÉTODO EN UN CASO REAL: SELECCIÓN DE MEZCLAS DE CCR PARA EL P.H. LAS CRUCES	55
4.1. Antecedentes	55
4.2. Expresión numérica del problema.....	57
4.2.1. Entradas.....	57

4.2.2. Salidas	60
4.3. Rangos dinámicos y distribución de las entradas y salidas	60
4.4. Gráficas globales.....	68
4.5. Análisis	77
4.6. Resultados obtenidos con minería de datos <i>versus</i> Resultados determinados en laboratorio	84
4.7. Ejemplos.....	88
CONCLUSIONES.....	98
REFERENCIAS	100

Índice de Tablas

Tabla 2-1. Requerimientos de la composición química de las puzolanas	22
Tabla 2-2. Granulometrías utilizadas en presas de CCR (tomada de ICOLD, 2003).....	25
Tabla 2-3 Coeficiente de Variación de varios tipos de ensayo en relación al nivel de control de calidad observado.....	30
Tabla 2-4. Contenidos de agua aproximados para diferentes t.m.a.....	33
Tabla 2-5. Volumen absoluto de agregado grueso por unidad de volumen de concreto para diferentes t.m.a.....	34
Tabla 4-1 Parámetros para cada zona de la cortina.....	89
Tabla 4-2 Resumen del ejemplo 1.....	92
Tabla 4-3 $E_{40\%}$ en función de los %'s de fracción V y finos	95
Tabla 4-4 Resumen del ejemplo 2.....	97

Índice de Figuras

Figura 1-1 Estructuras que conforman una presa	10
Figura 1-3. Secciones sin flujo y con el vertedor integrado para presas de gravedad	13
Figura 1-4. Sección típica de una cortina de CCR	14
Figura 1-5 Opciones de acabado en la cara aguas abajo	14
Figura 1-6 Opciones de acabado en la cara aguas arriba.....	15
Figura 1-7 Relación costo del CCR comparado con el Volumen de CCR colocado	16
Figura 1-8. Comparación de la sección de una cortina de enrocamiento contra una cortina de CCR.....	17
Figura 2-1. Materiales cementantes utilizados en presas de CCR.....	20
Figura 2-2. Cementos utilizados en presas de CCR.	21
Figura 2-3 Número de tamaños o fracciones de agregado utilizados en presas de CCR..	26
Figura 2-4 Permeabilidad vs. Contenido cementicio	28
Figura 2-5 Densidad in-situ vs Relación pasta/mortero para mezclas de CCR.....	29
Figura 3-1 Un sistema de Inteligencia Artificial IA	37
Figura 3-2. Ejemplo de un ajuste mediante una función parabólica.....	39
Figura 3-3. Un conjunto de entrenamiento o base de datos que determina completamente una función linealmente separable	41
Figura 3-4. Esquema de árbol de decisión.....	43
Figura 3-5. Un árbol de decisión implementando una función DNF	44
Figura 3-6. Un árbol de decisión implementando una lista de decisión	44
Figura 3-7. Árbol de decisión con atributos categóricos.....	45
Figura 3-8. Ocho patrones a ser clasificados mediante un árbol de decisión.....	47
Figura 3-9. Árbol de decisión univariado y su red equivalente	49
Figura 3-10. Árbol de decisión multivariado y su red equivalente.	49
Figura 3-11. Determinando cuando comienza el sobreajuste	51
Figura 3-12. Un árbol de decisión con subárboles replicados.....	53
Figura 3-13. Gráfico de decisión.....	53
Figura 3-14. Árbol de decisión multivariado.	54
Figura 4-1. Localización del proyecto	55
Figura 4-2. Sección y zonificación de la cortina	56
Figura 4-3 Escala de color en función de la relación agua/contenido cementicio	68

Figura 4-4. Regla de clasificación para f_c C/A.....	78
Figura 4-5. Árbol de decisión para f_c C/A.....	78
Figura 4-6. Regla de clasificación para f_c 28d.....	79
Figura 4-7. Árbol de decisión para f_c 28D	80
Figura 4-8. Reglas de clasificación para $E_{40\%}$ C/A.....	81
Figura 4-9 Diagrama en base a las reglas de clasificación obtenidas para $E_{40\%}$ C/A	82
Figura 4-10 Regla de clasificación para $E_{40\%}$ a 28d.	82
Figura 4-11 Regla de clasificación para f_t en C/A	83
Figura 4-12. Árbol de decisión para f_t en C/A.....	83
Figura 4-13 Resultados y rango de error para f_c C/A.....	84
Figura 4-14 Resultados y rango de error para f_c 28d.....	85
Figura 4-15 Resultados y rango de error para $E_{40\%}$ C/A.....	86
Figura 4-16 Resultados y rango de error para $E_{40\%}$ 28d.....	87
Figura 4-17 Resultados y rango de error para f_t C/A.....	88
Figura 4-18. Árbol de decisión basado en las reglas de clasificación obtenidas para $E_{40\%}$ C/A.....	90
Figura 4-19. Gráfica resistencia a compresión vs contenido cementicio para ejemplo 1...91	
Figura 4-20. Gráfica resistencia a compresión vs contenido cementicio para ejemplo 2...94	
Figura 4-21. Árbol de decisión basado en las reglas de clasificación obtenidas para $E_{40\%}$ (C/A)	95
Figura 4-22 $E_{40\%}$ en función del % de la fracción V y finos agregado a la mezcla	96

INTRODUCCIÓN

Durante todas las etapas que componen un Proyecto Hidroeléctrico PH, sobre todo cuando éste es de grandes dimensiones, se realizan una extensa variedad de estudios. Estas investigaciones están dirigidas a determinar y prospectar desde el impacto ambiental y las condiciones geológicas-geotécnicas, hasta las características detalladas de la obra de contención, de la obra de desvío, de la obra de excedencias, de la obra de toma, etc. Entre los análisis más importantes se encuentran los que tienen como objetivo evaluar las condiciones y propiedades de los materiales con los que se construirán las diferentes estructuras del Proyecto, principalmente las de aquellos relacionados con la obra de contención. A este respecto y sobre el caso de estudio abordado en este trabajo de tesis, el PH Las Cruces, a continuación se presenta una alternativa metodológica para estudiar y definir las mezclas de Concreto Compactado con Rodillo CCR, material del que se construirá la cortina.

Es interés particular para la Comisión Federal de Electricidad, ente que desarrolla el Proyecto citado, determinar a partir del número más eficiente posible de pruebas de laboratorio el proporcionamiento de mezclas de CCR con el que se consiguen las propiedades que se estipulan en los documentos que arrojó el diseño. Con base en con los comportamientos observados en el laboratorio, el personal de la Comisión debe establecer las condiciones de las mezclas bajo esquemas específicos que cumplan con los requerimientos del diseño de la obra; la solución a este punto no es trivial y exige modelado de alta dimensión -un gran número de variables reaccionando simultáneamente-. Hablando de mezclas de concreto, los componentes son de varios tipos y pueden ser utilizados en rangos propios muy particulares al sitio, a la economía del proyecto incluso a situaciones legales o sociales. En este contexto, el número de combinaciones de mezclas de CCR (de actuación condicionada y/o sincrónica) que se genera es notable. Encontrar experimentalmente todos los patrones de comportamiento (mecánico), el grado en el que influyen los ingredientes y sus diferencias en estado fresco y endurecido, es una labor muy difícil, económicamente prohibitiva y bajo ciertos enfoques científicos, pudiera resultar engañosa.

En este trabajo de investigación se propone, para cumplir con los objetivos de diseño y de construcción de manera efectiva, el uso de técnicas de la Minería de Datos para optimizar el proceso de selección de mezclas. Se presenta la aplicación de algoritmos MD que generan Árboles de Regresión y se explica la manera en la que, a través de un reducido conjunto de pruebas, un ingeniero constructor puede maximizar la información que se genera experimentalmente con árboles que permiten simular comportamientos de las

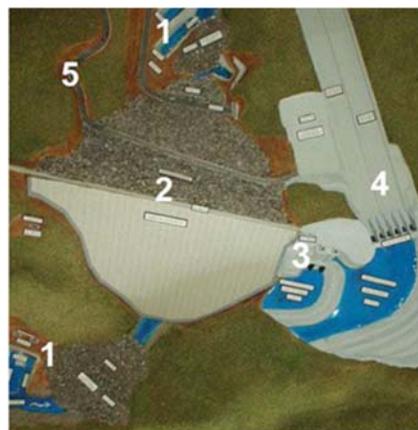
mezclas de CCR más allá de las fronteras marcadas por las propias pruebas. Ésta práctica es esencialmente ventajosa pues permite, además de una eficiente planeación de obra, modificaciones o ajustes por carencias o deficiencias en materiales y correcciones *a posteriori* (después de diagnósticos *in situ*) de las dosificaciones en cada zona de la cortina, sin que necesariamente se repita o amplíe el programa de ensayos.

Con los resultados obtenidos, más sólidos científicamente y más sonoros empíricamente, se pueden determinar i) las mezclas con las que se consiguen las demandas de diseño (resistencia a la compresión f_c , módulo elástico E y resistencia a la tensión indirecta f_t), ii) las directrices de futuros estudios de CCR (más económicos, más específicos, más ambiciosos) y; iii) los aspectos que hacen crecer el estado del arte y el estado de la práctica y que podrían conformar nuevas líneas de investigación.

1. CAPÍTULO I. CONCRETO COMPACTADO CON RODILLO (CCR) EN PRESAS DE GRAN ALTURA.

Las grandes presas son un fuerte icono de modernización, prestigio nacional y dominación humana sobre la naturaleza. Son barreras artificiales colocadas sobre el cauce de un río con el propósito de retener el agua que fluye por el cauce para después utilizarla en propósitos sociales estratégicos. La obra de contención es solo una parte de todo un sistema (*Figura 1-1*) en el que existen estructuras permanentes (como la cortina misma, obra de toma, obra de excedencias, caminos de acceso y operación, caseta de operación, casa de máquinas) y estructuras temporales (como la obra de desvío, caminos de construcción y campamentos) que en conjunto con una serie de servicios (teléfono, energía eléctrica, estación climatológica, etc.) permiten aprovechar adecuadamente e l agua almacenada.

<i>Id</i>	<i>Estructura</i>
1	Obra de desvío
2	Obra de contención
3	Obra de toma
4	Obra de excedencias
5	Caminos de acceso

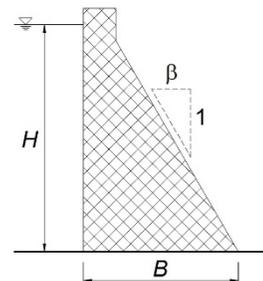


*Figura 1-1 Estructuras que conforman una presa
(Tomado de Pérez, 2013)*

Clasificación de las presas

Las presas se pueden clasificar en diferentes categorías, entre las maneras más empleadas para categorizarlas están (Valencia, 2006 y, Pérez; 2013):

- a) Según el proyecto hidráulico:
- Cortinas vertedoras
 - Cortinas no vertedoras
- b) Según los materiales de construcción:
- Presas de Tierra
 - Presas de enrocamiento
 - Presas de materiales graduados
 - Presas de concreto
 - **Presas de concreto compactado con rodillo CCR**
- c) Según su uso:
- Presas de almacenamiento,
Tienen como objeto el abastecimiento de agua para recreación o cría de peces y animales, *generación eléctrica*, irrigación, etc.
 - Presas de derivación
Tienen como objeto el desvío del agua en sistemas de conducción para el consumo urbano o de procesos industriales.
 - Presas reguladoras
Tienen como objeto el control de avenidas y regulación de escurrimientos durante la temporada de sequías.
 - Combinación de las anteriores.
- d) Según la forma de trabajo estructural
- Rígidas: generalmente son de concreto
 - Flexibles: son de rellenos de suelos y/o enrocado, con taludes muy tendidos.
- e) Según la disposición de la cortina en planta
- De eje recto
 - De eje curvo
 - De eje mixto
- f) Según la esbeltez de la cortina (β)
- Presas flexibles $\beta \geq 1.0$
 - Presas de gravedad $0.6 \leq \beta < 1.0$
 - Presas de arco-gravedad $0.3 \leq \beta < 0.6$
 - Presas de arco puro $\beta < 0.3$



Factores que determinan el tipo de cortina

- i) Topografía de la boquilla.
- En general, se consideran tres tipos de boquillas: a) en forma de "V"; b) en forma de "U" o en cañón y; c) abierta o alargada.

- ii) Geología de la boquilla.
El tipo de roca o suelo en el eje donde se cimentará la cortina, sus propiedades mecánicas y la presencia de discontinuidades son determinantes en la selección del tipo de cortina.
- iii) Disponibilidad de materiales.
Este factor es determinante para la selección del tipo de cortina ya que incide directamente en el costo de la obra, específicamente de los acarreos.
- iv) Magnitud de la cortina.
Impacta directamente en las cargas, esfuerzos de trabajo de diseño y en volúmenes de materiales.
- v) Efecto sísmico.
La zona sísmica donde se ubique el proyecto puede ser decisiva para la selección del tipo de cortina.
- vi) Clima del lugar.
Periodos de retorno de avenidas, temporada de estiaje, afectación de materiales de construcción, son aspectos que se asocian a este factor.
- vii) Aspectos Viales.
Las zonas inundadas por el embalse podrían provocar la reubicación de caminos y vías de comunicación existentes.

Presas de gravedad de CCR

Las presas de gravedad son estructuras de concreto sólido que mantienen su estabilidad contra las cargas de diseño gracias a su forma geométrica y a la masa y resistencia del concreto. Generalmente están construidas con un eje recto aunque éste puede ser ligeramente curvo o angular para adecuarse a las condiciones de un sitio específico. Las presas de gravedad comúnmente consisten de una sección sin flujo y una con el vertedor integrado como se muestra en la *Figura 1-2*.

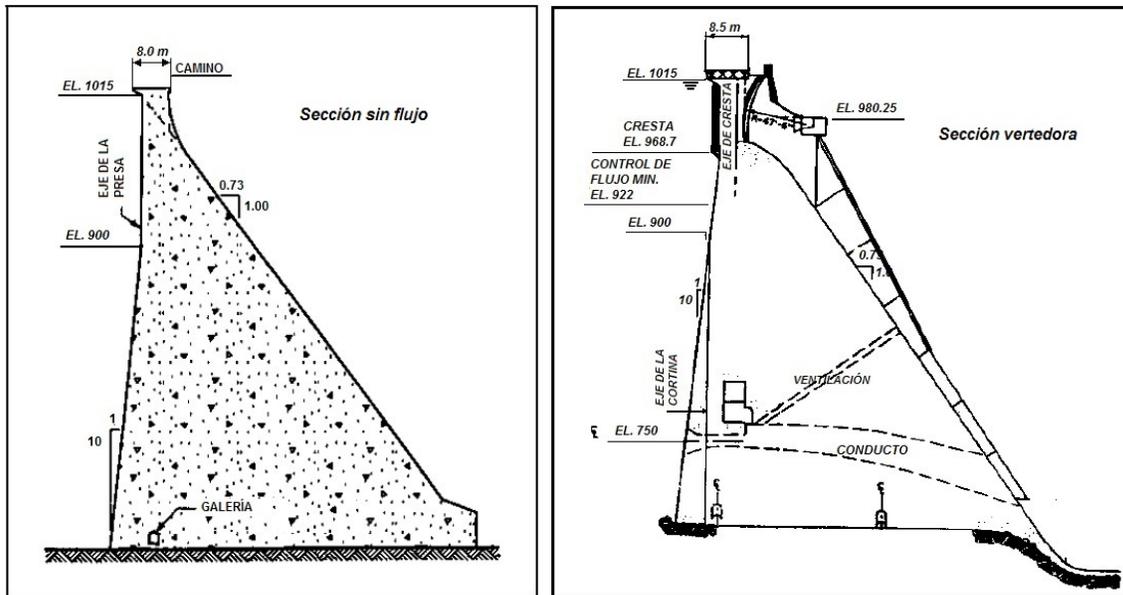


Figura 1-2. Secciones sin flujo y con el vertedor integrado para presas de gravedad (Modificado de USACE, 1995)

Los dos métodos de construcción más utilizados para presas de gravedad son el concreto convencional masivo y el concreto compactado con rodillo CCR. El presente trabajo está enfocado al estudio del CCR en la construcción de presas sin embargo se utilizarán algunas descripciones sobre concreto convencional masivo para dirigir los comparativos y fundamentar las ventajas de su uso.

Las presas de gravedad construidas bajo el método de CCR resultan más económicas, comparadas con las realizadas con concreto convencional debido, fundamentalmente, a las técnicas de colocación rápida. Los procedimientos de construcción asociados al CCR requieren atención particular en la distribución y el diseño de la estanqueidad y control de filtraciones, juntas horizontales y de contracción, elementos de las caras aguas arriba y aguas abajo y estructuras complementarias (ver *Figura 1-3*, *Figura 1-4* y *Figura 1-5*). Un diseñador capacitado aprovecha las ventajas que concede el CCR y usa estratégicamente los conocidos criterios de ingeniería de grandes presas para obtener el equilibrio entre reducción de costos y esfuerzos técnicos relacionados con la seguridad, la durabilidad, y el rendimiento a largo plazo (USACE, 1995).

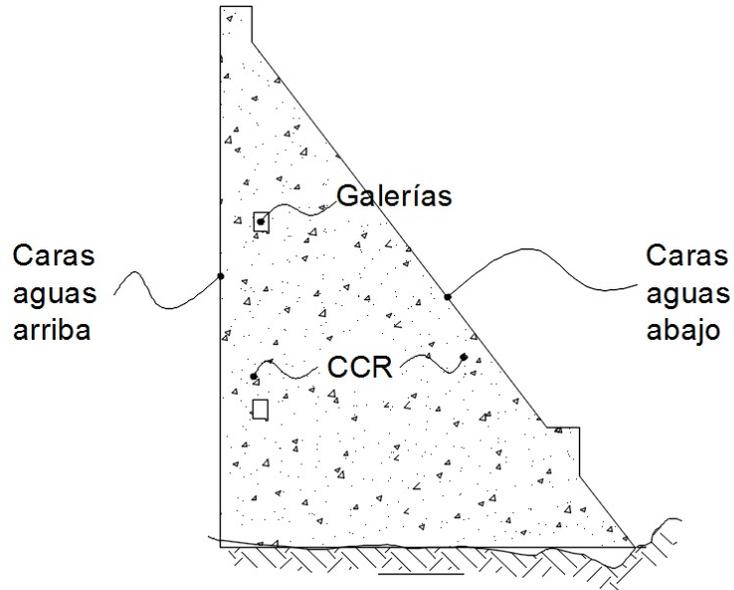


Figura 1-3. Sección típica de una cortina de CCR

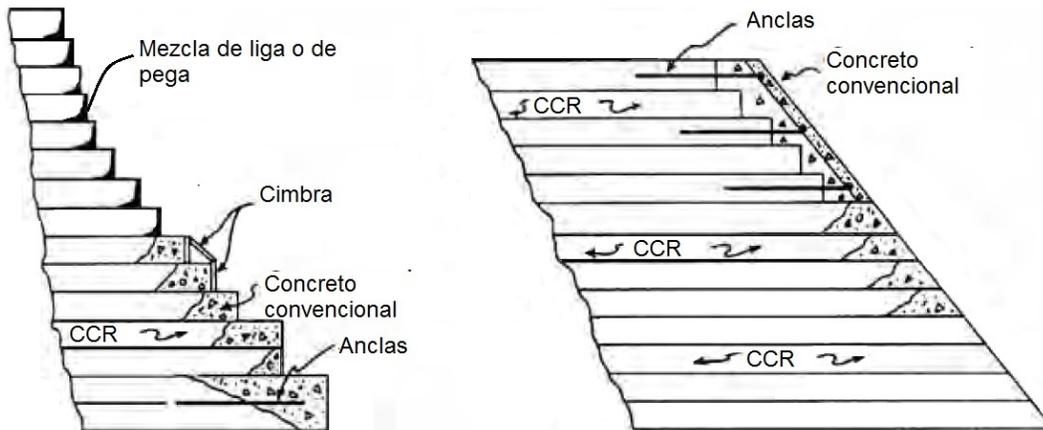


Figura 1-4 Opciones de acabado en la cara aguas abajo
(Modificado de Nawy, 1998)

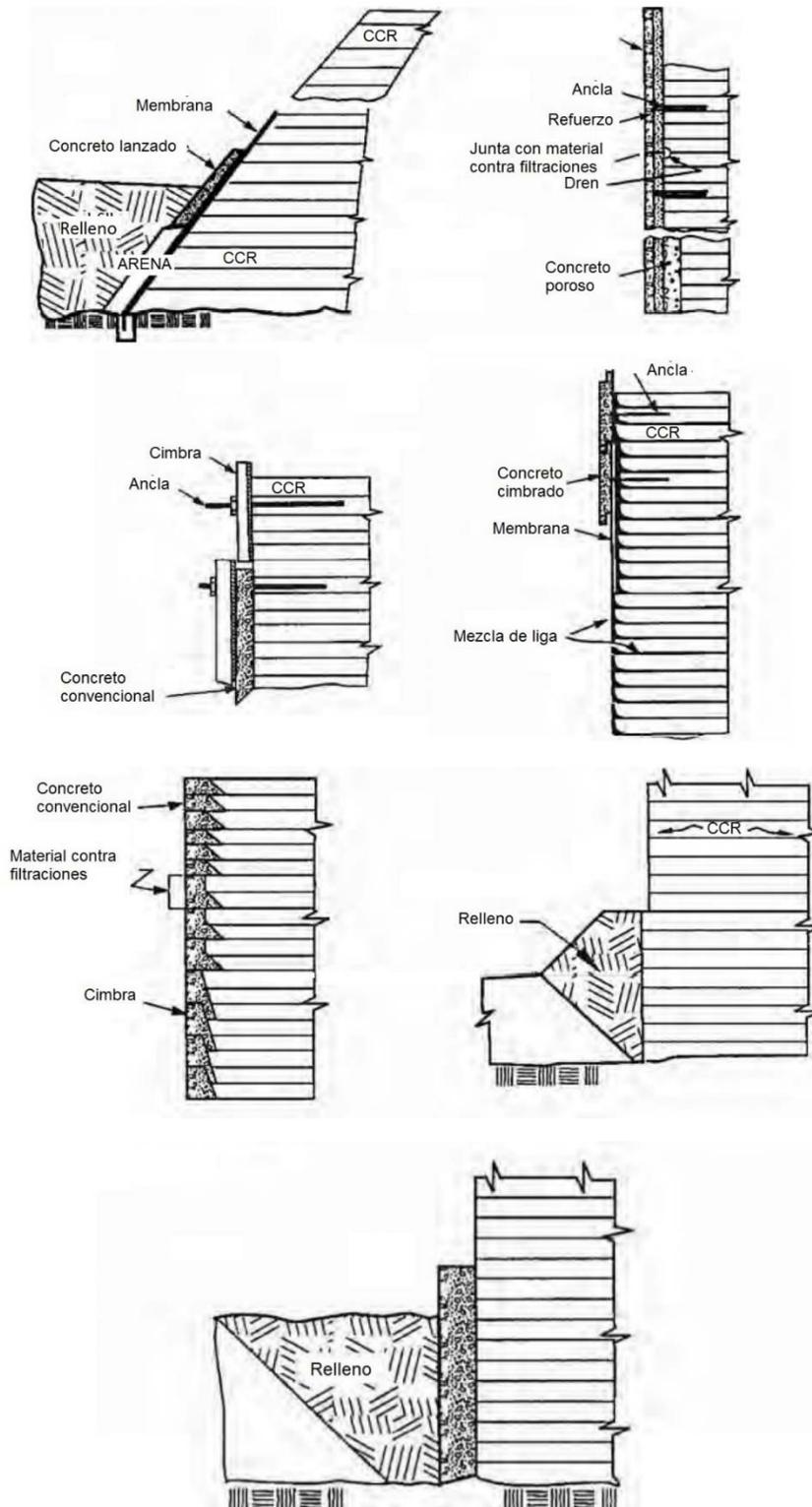


Figura 1-5 Opciones de acabado en la cara aguas arriba
(Modificado de Nawy, 1998)

Ventajas del CCR en la construcción de grandes presas

La construcción de presas de CCR tiene aspectos positivos relevantes y delicadas particularidades que deben ser tomadas en cuenta para asegurar proyectos exitosos. Mientras que las ventajas están relacionadas con tipos de mezclas, diseños estructurales, métodos de producción, aptitud contra clima, las dificultades derivan de los retos y condiciones propias de la ejecución de la obra en sitios particulares (ACI 207.5,1999).

Algunos de los aspectos por los que la construcción con CCR es altamente competitiva son (USACE, 2000):

- a) Costos. El costo por unidad de volumen del CCR es de un 25 y hasta 50% menor que el costo del concreto convencional. La diferencia en el porcentaje economizado depende del costo del agregado y del material cementante, de la complejidad de la colocación, y de la cantidad total de concreto colocado. Los ahorros asociados con el CCR comparado con concreto convencional se deben principalmente a la reducción de cimbra, tiempos de colocación, y costos de compactación.

La Figura 1-6 muestra la relación entre el costo del CCR comparado con el volumen de CCR colocado basado en proyectos construidos en los Estados Unidos.

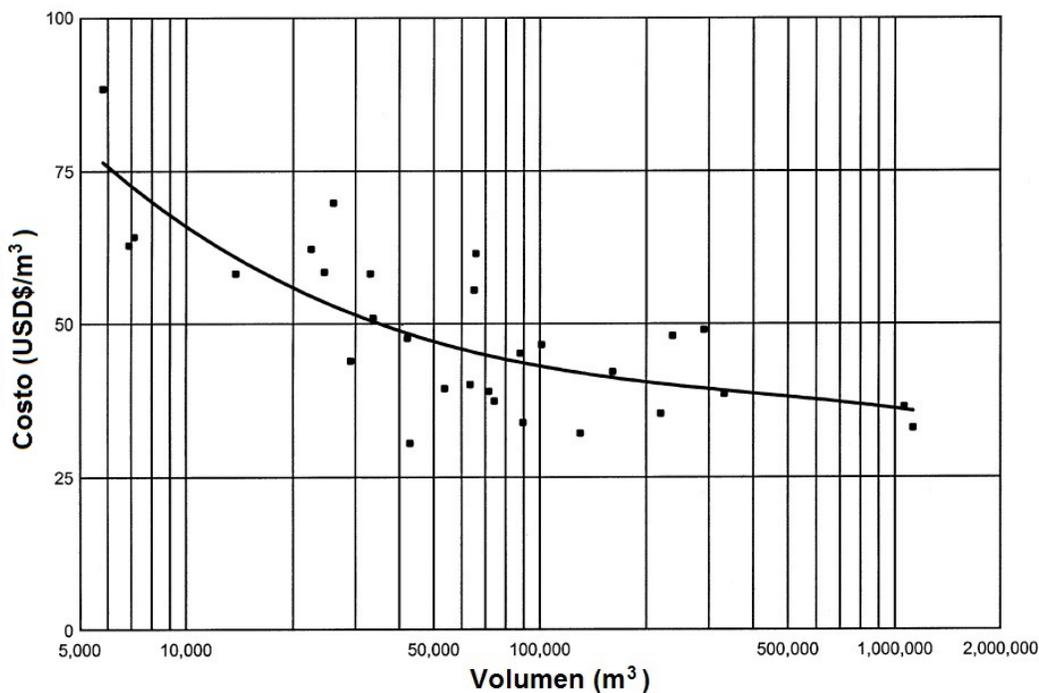


Figura 1-6 Relación costo del CCR comparado con el Volumen de CCR colocado (proyectos hechos en Estados Unidos, tomado de USACE, 2000)

- b) Construcción rápida. El empleo de técnicas de construcción rápida (velocidad comparada con aquellas desarrolladas cuando se construyen presas de concreto convencional y de enrocamiento con cara de concreto ECC) y la reducción del volumen de material empleado, tienen un gran impacto en la disminución de costos en presas de CCR. El proceso constructivo con CCR considera a la

colocación del material como una actividad casi continua lo que posibilita la obtención de altos rendimientos de producción (recorta el periodo de construcción de la presa). Otros beneficios derivados de una rápida construcción incluyen la disminución de costos administrativos, la obtención pronta de los beneficios del proyecto, y la posible modificación o incluso la eliminación de la obra de desvío.

- c) Vertedores y obras complementarias integrales. De la misma forma en que se diseña en presas de concreto convencional, para las presas de CCR el vertedor puede ser incorporado dentro de la estructura de contención. El arreglo típico permite la descarga sobre la cresta de la presa y fluye sobre la cara aguas abajo. Para proyectos que requieren una obra de toma "multinivel" para el control de calidad del agua o para prevenir depósitos de sedimentos (azolves), la obra de toma puede ser anclada fácilmente a la cara aguas arriba de la presa de CCR.

El costo de una obra de toma en una presa de CCR es considerablemente menor que el costo de la obra de toma en una presa de enrocamiento, debido a la menor longitud de los túneles de conducción de la obra de toma. En la *Figura 1-7*, se hace una comparación entre secciones de una presa de enrocamiento y otra de CCR, correspondientes a una presa de 100 m de altura; se puede observar que la longitud de la base es más corta para una presa de CCR que para una presa de enrocamiento; esto reduce la longitud del conducto y el tamaño de las compuertas de descarga, la longitud de la obra de desvío y el tamaño de las obras para generación de energía; además del importante decremento en el costo de la preparación de la cimentación.

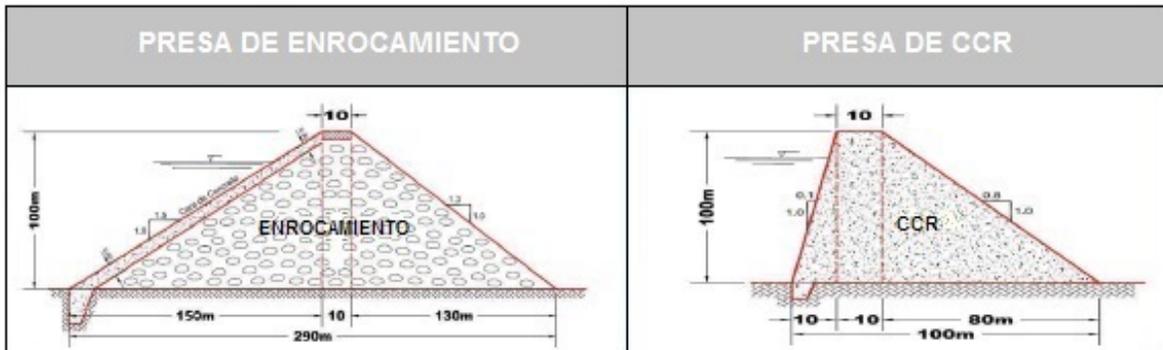


Figura 1-7. Comparación de la sección de una cortina de enrocamiento contra una cortina de CCR

- d) Desvíos y ataguías estratégicos. Las presas de CCR proveen ventajas en el desvío estratégico del río y permiten reducir los daños y riesgos asociados con el desbordamiento de las ataguías. Los túneles de desvío para presas de CCR son más cortos que para las presas de enrocamiento, tienen periodos reducidos de construcción, y la probabilidad de grandes avenidas es menor (con ello las dimensiones de la obra de desvío y la altura de las ataguías son menores a las requeridas para presas de enrocamiento o concreto convencional).
- e) Resistencia a la erosión y desbordamiento. La resistencia a la erosión del CCR ha sido sorprendentemente buena, (Schrader y Stefanakos, 1995) han publicado un resumen histórico de éste parámetro. Evaluaciones recientes, incluyendo pruebas

a gran escala con velocidades de 30 m/s, han sido documentadas por el U.S. Army Corps of Engineers.

Para mezclas de buena calidad, con contenidos cementicios del orden de 100-180 kg/m³, se obtuvieron rangos de erosión de 10g y 244g por m² en superficie lisa y rugosa respectivamente. La experiencia en campo confirma que esos valores son razonable y conservadores (Nawy, 2008). En otro estudio (Wang Xin *et al*, 2012) se realizaron un estudio para determinar los rangos de erosión en laboratorio en losas de CCR de 15x15x30 cm con velocidades de flujo de 30-35 m/s. Obteniendo resultados satisfactorios. Ambas referencias muestran al CCR como un material resistente a la erosión. Por el contrario, la principal causa de la falla de una presa de tierra y enrocamiento es por desbordamiento de la cortina; y la segunda causa de falla corresponde a la tubificación del cuerpo de la cortina (Murillo, 2012). Ambos casos están relacionados con la erosión de la cortina. El flujo de agua proveniente del embalse, ya sea por encima de la corona o a través del cuerpo de la cortina y/o cimentación, es la fuerza que arrastra el material provocando la inestabilidad de la cortina y su consecuente falla. Conforme aumenta la erosión, aumenta el flujo de agua y en consecuencia aumenta la erosión.

Por lo mencionado en este inciso, se puede decir que una cortina de CCR es más resistente a la erosión y al desbordamiento que una cortina de enrocamiento.

Otros aspectos que deben ser tomados en cuenta es que a menores volúmenes de materiales la localización de los bancos de agregados es un factor menos determinante en la selección del sitio, además, el banco de préstamo podrá ser de menor escala y tendrá mayor aceptación desde el punto de vista ambiental.

El concreto compactado con rodillo también puede ser utilizado en presas con otros propósitos (ICOLD, 2003), como son:

- el refuerzo y ampliación de una presa ya existente;
- el refuerzo de la cara aguas abajo de presas de materiales sueltos de tal forma que puedan tener vertido por coronación;
- cimentaciones
- el relleno de depresiones en presas de materiales sueltos o de cimentaciones debidas al vertido por vertedor o por coronación, etc.
- apoyo de estructuras auxiliares como conducciones de desagües, etc.

2. CAPÍTULO II. CARACTERÍSTICAS DEL CCR

La rapidez de la puesta en obra, el relativamente bajo contenido de cemento y la utilización de adiciones minerales (cenizas volantes, “filler” calizo, residuos mineros, etc.) son motivos suficientes para calificar a este material como económicamente atractivo para la industria de la construcción y por lo tanto los esfuerzos en investigación están puestos en desarrollar métodos que hagan más eficientes el diseño (de mezclas), los procesos de colocación, y el monitoreo del comportamiento (Escalaya, 2001). En este trabajo de tesis se aborda únicamente el primero de estos aspectos: el diseño de mezclas.

El CCR es un material de consistencia seca con revenimiento cero que en estado fresco puede soportar el tránsito de maquinaria pesada como tractores y rodillos (equipo de distribución y compactación del CCR) al tiempo que tiene la humedad suficiente para permitir la adecuada distribución del mortero durante el mezclado y la compactación vibratoria.

En la planta de concreto, el CCR es fabricado y transportado por equipos que trabajan continuamente y es acarreado mediante camiones o, de manera más eficiente, en bandas transportadoras. El tendido del CCR se ejecuta en capas de espesores constantes (recomendación práctica en cortinas como la elegida en el proyecto hidroeléctrico estudiado en este trabajo de tesis es ~30 cm).

2.1. Materiales

La consistencia del CCR debe ser tal que pueda ser compactada mediante rodillos vibratorios. La granulometría de los agregados y el contenido de pasta son elementos cruciales en el diseño de la mezcla. Por sus particulares condiciones, para el CCR se han desarrollado métodos específicos que permiten estimar sus propiedades en estado fresco y que difieren de los utilizados para conocer estas mismas características en el concreto convencional.

Debe tomarse en cuenta que las evaluaciones del CCR se generan para su colocación de manera masiva y que las características del CCR en sitio están basadas en el enfoque del diseñador, los criterios de diseño, y el control en la producción del mismo. Así pues, la selección de materiales y el criterio para proporcionamiento que se emplee al fabricar las mezclas son tan importantes como las pruebas de laboratorio para evaluar las propiedades en estado fresco y endurecido del mismo. Si bien existen recomendaciones para la selección de materiales (por ejemplo las provistas en *ACI 207.1R “Guide to Mass Concrete”*) es el conocimiento y experiencia del diseñador sumados a los retos que imponga el sitio del proyecto lo que determina las características del material empleado.

Los materiales se pueden clasificar en i) cementantes, ii) agregados pétreos y iii) aditivos químicos.

2.1.1. Materiales cementantes

El CCR puede ser elaborado con cualquiera de los tipos básicos de cemento o más comúnmente con una combinación de cemento y adiciones minerales.

Los diferentes materiales cementantes que han sido utilizados en las presas de CCR, terminadas hasta finales de 1996, se muestran en la *Figura 2-1*. Como se observa en el gráfico la gran mayoría de las mezclas de CCR contienen adiciones minerales, en particular ceniza volante de bajo contenido de cal. Sólo el 10.2% de las mezclas de CCR incluyen únicamente cemento Portland mientras que el 4.5% contiene una combinación de adiciones minerales (por lo regular escoria granulada con ceniza volante con alto contenido de cal) y polvo triturado sin cemento.

El uso de adiciones minerales tiene como objetivo minimizar el contenido de cemento Portland, con lo que se logra un menor costo y una reducción del calor de hidratación (esto último provoca un lento desarrollo de la resistencia con lo cual se mejora la integración entre capas de CCR y se reducen los esfuerzos térmicos).

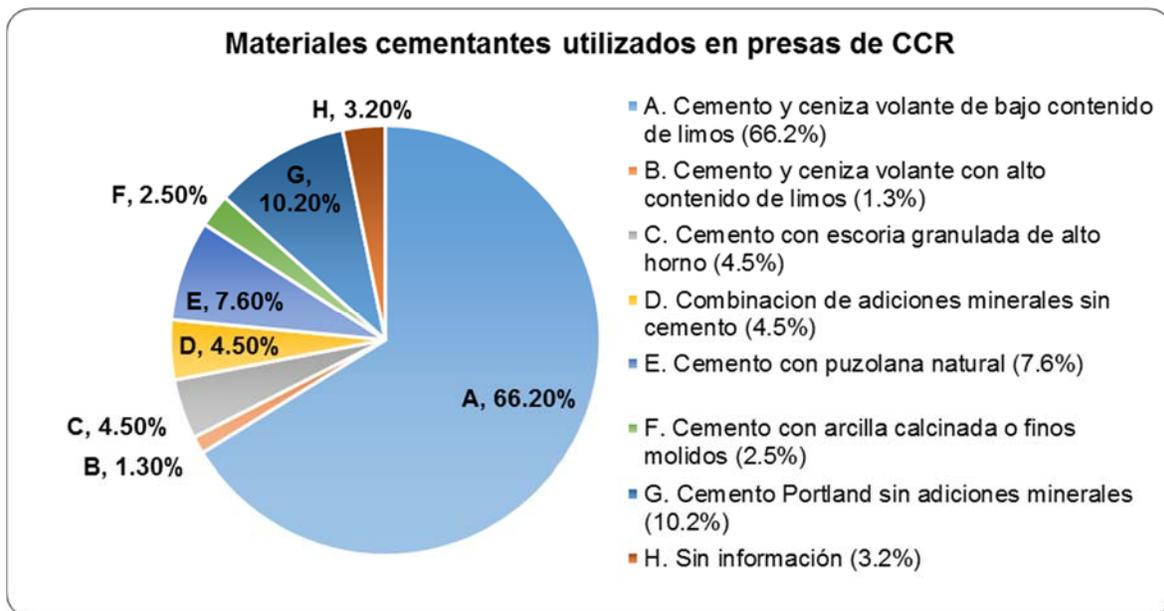
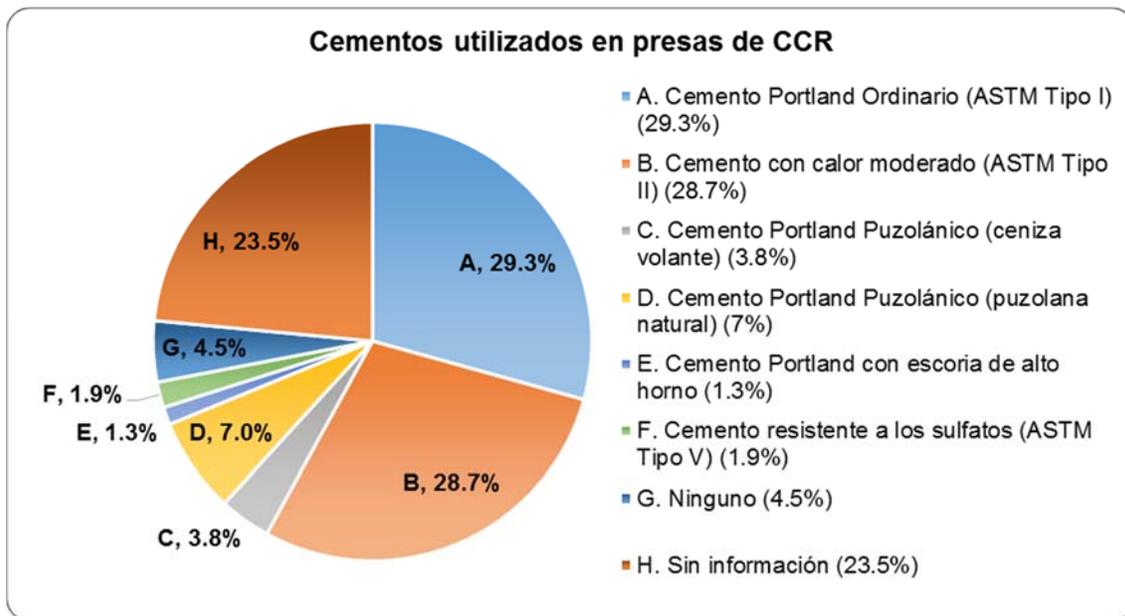


Figura 2-1. Materiales cementantes utilizados en presas de CCR (Tomado de ICOLD, 2003)

a) Cemento

El CCR se puede elaborar con cualquiera de los tipos comunes de cemento, sin embargo para presas de CCR es preferible utilizar cementos de menor calor de hidratación que el cemento Portland Ordinario CPO (Tipo I, ASTM C 150). Entre los recomendados se encuentran los Tipo II (moderado calor), Tipo IP (cemento Portland Puzolánico), y Tipo IS (cemento Portland con escoria granulada de alto horno) *Figura 2-2*. El desarrollo de la resistencia para estos cementos de bajo calor de hidratación es más lento que para el cemento Portland Ordinario a edades tempranas, sin embargo a edades tardías los cementos que tienen un desarrollo lento por lo regular producen resistencias mayores que las del cemento Portland Ordinario.



*Figura 2-2. Cementos utilizados en presas de CCR.
(Tomado de ICOLD, 2003)*

Aunque la efectividad de los cementos bajo control de temperatura ha sido probada, una alternativa viable puede ser utilizar cemento Portland Ordinario con adiciones minerales para lograr los mismos efectos. Esta solución en ocasiones es más económica, sobre todo en aquellos sitios donde los cementos de bajo calor están menos disponibles (y por ende son generalmente más costosos) que el cemento Portland ordinario.

b) Adiciones minerales

Son materiales silíceos finamente molidos que se añaden al cemento. Algunas adiciones minerales son puzolánicos (por ejemplo: ceniza volante de bajo contenido de cal, puzolanas naturales o arcilla calcinada), otros son cementicios (por ejemplo escoria granulada de alto horno), mientras que otros son tanto cementantes como puzolánicos (ceniza volante de alto contenido de cal).

La ASTM C 618, maneja tres clasificaciones de puzolanas:

- I. Clase N. Puzolanas naturales sin calcinar o calcinadas que cumplan con los requerimientos mostrados en la *Tabla 2-1* Pueden ser tierras diatomáceas, pizarras de sílex, tobas y cenizas volcánicas o pumíticas (calcinadas o sin calcinar) y diferentes materiales que requieren calcinación para inducir propiedades satisfactorias, tales como algunas arcillas y lutitas.

Tabla 2-1. Requerimientos de la composición química de las puzolanas (tomada de ASTM C 618)

	Clase		
	N	F	C
SiO ₂ + Al ₂ O ₃ + Fe ₂ O ₃	70.0	70.0	50.0
SO ₃ , max., %	4.0	5.0	5.0
Contenido de humedad, max., %	3.0	3.0	3.0
Pérdida por calcinación, max., %	10.0	6.0*	6.0

* El uso de puzolana clase F con una pérdida por calcinación mayor al 12.0% puede ser aprobada por el usuario si tanto los registros de desempeño como los resultados de laboratorio son aceptables.

- II. Clase F. Ceniza volante resultado de la quema de antracita o carbón bituminoso, y debe cumplir con los requerimientos solicitados en el documento ASTM C 618. Esta clase de ceniza volante tiene propiedades puzolánicas.
- III. Clase C. Ceniza volante normalmente producida a partir de carbón sub-bituminoso. Este tipo de ceniza volante tiene propiedades puzolánicas y algunas propiedades cementicias.

Las puzolanas que cumplen con las especificaciones de la norma ASTM C 618 para las clases C, F y N, se han utilizado con éxito en mezclas de CCR. Las puzolanas Clase F y N son las más utilizadas, ya que normalmente contribuyen a generar menor calor de hidratación que la Clase C y tienen mayor resistencia a los sulfatos.

La escoria granulada es más parecida al cemento Portland que la ceniza volante o las puzolanas naturales sobre todo en sus propiedades y en que tiene un tiempo de fraguado más rápido y un calor de hidratación relativamente alto comparado con otras adiciones minerales. Los componentes de la escoria (cal, sílice y alumina) son los mismos que para el cemento, pero en diferentes proporciones.

El uso de las adiciones minerales dependerá del rendimiento de material requerido así como del costo y la disponibilidad en cada proyecto; su uso en las mezclas de CCR puede servir a uno o más de los siguientes propósitos:

- como una porción del contenido cementicio para reducir la generación de calor,
- como una porción del contenido cementicio para reducir el costo y,
- como una adición para proveer finos suplementarios para mejorar la trabajabilidad de la mezcla y para formar parte del volumen de pasta.

Las cenizas volantes de bajo contenido de cal, particularmente cuando se utilizan como una porción significativa del material cementante, pueden retrasar el fraguado inicial de las mezclas de CCR, contribuyendo a extender el tiempo de vida de la junta “caliente” entre capas. El incremento de calor en mezclas de CCR con alto contenido de adiciones minerales puede continuar de 60 a 90 días después de la colocación y tal incremento se debe considerar en cualquier estudio térmico.

La actividad puzolánica continúa después de varios años, lo que indica que las puzolanas pueden ocupar espacio y también pueden contribuir al desarrollo de la resistencia a largo plazo.

Es importante señalar que la norma citada en esta sección, de uso amplio a nivel mundial, maneja algunas propiedades químicas y físicas que las adiciones deben cumplir para obtener alguna de las clasificaciones y aquel que las usa debe cumplir cabalmente estos requerimientos para asegurar la cota mínima en la calidad de producción del concreto.

2.1.2. Agregados

La selección y control de la granulometría de los agregados son factores que influyen en la calidad *in situ* y en las propiedades del CCR. La variación de los agregados durante la construcción afectará significativamente los requerimientos de agua y cemento, y en consecuencia afectaran la resistencia y el rendimiento.

Agregados similares a los utilizados en el concreto convencional se han utilizado en el CCR, incluso agregados que no cumplen los estándares requeridos para concreto convencional han sido utilizados satisfactoriamente en la construcción de CCR. La mayoría de los proyectos de CCR se han construido con agregados que cumplen los requisitos de la ASTM C 33, sin embargo, existen excepciones interesantes y exitosas como las del aumento en la cantidad de finos que pasan por la malla no. 200 (0.075mm) para obtener mejores propiedades mecánicas.

Los rendimientos de colocación del CCR son por lo general altos, por lo tanto es recomendable tener grandes reservas de agregados para prevenir que los rendimientos de colocación dependan de la producción de agregado. Esto puede propiciar que se empleen agregados de relativa baja calidad pero en los que la ubicación del banco sea tan cercana al sitio que dirija al proyecto hacia niveles más económicos.

Un objetivo básico en el proporcionamiento de cualquier concreto es incorporar la máxima cantidad de agregado y la mínima cantidad de agua a la mezcla, de este modo reducir el contenido de cemento, y en consecuencia reducir los cambios de volumen del concreto. Este objetivo se logra con una adecuada granulometría, al definir la granulometría, también se debe tener cuidado en escoger un t.m.a. (tamaño máximo de agregados) que evite o reduzca la segregación del mismo. La experiencia ha mostrado que esto se ha logrado con t.m.a. de 50 mm (2”) en los que la cantidad de material que pase la malla no. 4 (4.75 mm) sea mayor a la utilizada para el concreto convencional (ACI 207.5, 1999).

La presencia de partículas planas y/o alargadas es tradicionalmente indeseable en la industria del concreto, sin embargo, las mezclas de CCR parecen ser menos afectadas por estas partículas que las mezclas de concreto convencional. El Cuerpo de Ingenieros de EU actualmente tiene como límite el 25% en el contenido permitido de partículas planas y alargadas y los resultados de comportamiento han sido satisfactorios.

Además de la necesidad de dureza, durabilidad del agregado, alto peso volumétrico, las características que afectan las propiedades térmicas y agrietamiento de la presa son importantes. Se buscan bajos módulos elásticos y bajo coeficiente de expansión térmica.

- a) Agregado grueso. La selección del t.m.a. debe basarse en la necesidad de reducir las cantidades de materiales cementicios, controlar la segregación y facilitar la compactación. La selección de la separación de los tamaños de agregado puede hacerse conforme al ACI 304R, aunque puede ser específica en cada proyecto.
- b) Agregado fino. La distribución del agregado fino influye fuertemente en los requerimientos de la pasta y de compactación del CCR. También afecta las cantidades de agua y material cementante requeridos para llenar los huecos.
- c) Finos. Los materiales finos son aquellos que pasan la malla no. 200 (0.075mm). La cantidad y características de los finos afectarán la compactación relativa de la mezcla de CCR y pueden influir en el número de pasadas del rodillo durante la construcción de la presa. Los finos llenan el espacio vacío entre los agregados, proporcionan una consistencia compactable, ayudan a controlar la segregación y disminuyen la permeabilidad. El exceso de finos es perjudicial al reducir la trabajabilidad y demandar mayor cantidad de agua (y las consecuencias adversas conocidas de esta situación).

Granulometría.

Los rangos típicos de granulometría utilizados en las presas de CCR se muestran en la *Tabla 2-2*, en dicha tabla, entre otros datos es interesante observar los casos extremos de las presas, la Jordão con un alto contenido de agregado fino, significativamente por encima del resto de presas de su tipo, y la Santa Eugenia, que tiene un contenido muy bajo de agregado fino, ambas con comportamientos satisfactorios. Normalmente las presas de CCR con contenido de cementante bajo tienen granulometrías con porcentajes más elevados de finos que aquellas con contenidos de cementante más altos. Hay una tendencia clara a la reducción del contenido del agregado fino con el incremento del tamaño máximo del agregado y con el incremento de trabajabilidad.

Tabla 2-2. Granulometrías utilizadas en presas de CCR (tomada de ICOLD, 2003)

	Zintel Canyon	Jordão	Shimajigawa	Miyagase	Copperfield	Petit Saut	De Mist Kraal	Upper Still-water	Santa Eugenia	New Victoria	Pangue	Platanovryssi
Clasificación	Bajo	Bajo	PHR	PHR	Medio	Medio	Medio	Alto	Alto 88	Alto	Alto	Alto
Cemento	74	65	84	91	80	0	58	79	152	79	80	50
Puzolana	0	10	36	39	30	130	58	173	Media	160	100	225
Trabajabilidad	Media	Baja	Alta	Alta	Baja	Baja	Media	Alta		Alta	Media	Alta
Tamaño de tamiz (mm)												
150				100								
100				82								
75			100	7			100		100			
62	100	100		74			91		94			
50	98	98		70	100	100	81	100	84		100	
38	91	89	76	65	90	90	67	95	74	100	99	100
25	77	78		58		73	54		61		86	
19	70		58	51	68	67	46	66	52	70	76	66
9.5	50	62	43	39	54	52	35	45	34	48	61	48
4.75	39	51	33	30	44	37	29	35	24	33	41	37
2.4	25	40	30	26	36	30	22	26	12	26	24	29
1.2	18	26	23	19	28	25	18	21	6	21	20	21
0.6	15	18	14	11	21	20	16	17	3	15	15	
0.3	12	15	8	6	15	14	14	10	2	7	6	8
0.15	11	10	3	2	10	11	6	2	1	3	2	
0.075	9	8	0	0	7	7	2	0	0	0	1	

Nota: Trabajabilidad baja - Tiempo VeBe no posible; Trabajabilidad media - Tiempo VeBe>25s; Trabajabilidad alta - Tiempo VeBe<25s

Un CCR con un tamaño máximo de agregado de 100 mm, alta trabajabilidad y bajo tiempo VeBe (el tiempo o prueba VeBe es una medida de la consistencia de las mezclas de concreto extremadamente secas), se puede esperar que tenga un contenido de agregado fino del 30 %, mientras que un CCR con un *t.m.a.* de 50 mm, trabajabilidad baja y tiempo VeBe alto, se puede esperar que tenga un contenido de agregado fino del 45 %.

Cuando se desea un control exhaustivo de la granulometría de los agregados y de la producción de CCR, las separaciones de tamaño deben seguir la práctica normal del concreto (Hollingworth, *et al*, 1986). Se pueden observar ahorros aparentes en el costo combinando dos o más rangos de tamaño para reducir el número de acopios, sin embargo, a medida que el rango del tamaño aumenta en cada acopio, aumenta también la dificultad en evitar la segregación de las partículas más grandes durante la manipulación del agregado. El proyectista y/o el contratista deben considerar los potenciales ahorros de costos en la reducción del número de acopios y separar la manipulación y las facilidades de pesado frente al potencial de incremento de variación en la granulometría del agregado y su impacto en la uniformidad de la consistencia y resistencia, en la unión entre juntas, y en la permeabilidad del CCR resultante.

La *Figura 2-3* muestra el número de tamaños o fracciones de agregado utilizados en la mayoría de presas de CCR (estadística para presas acabadas al final de 1996).

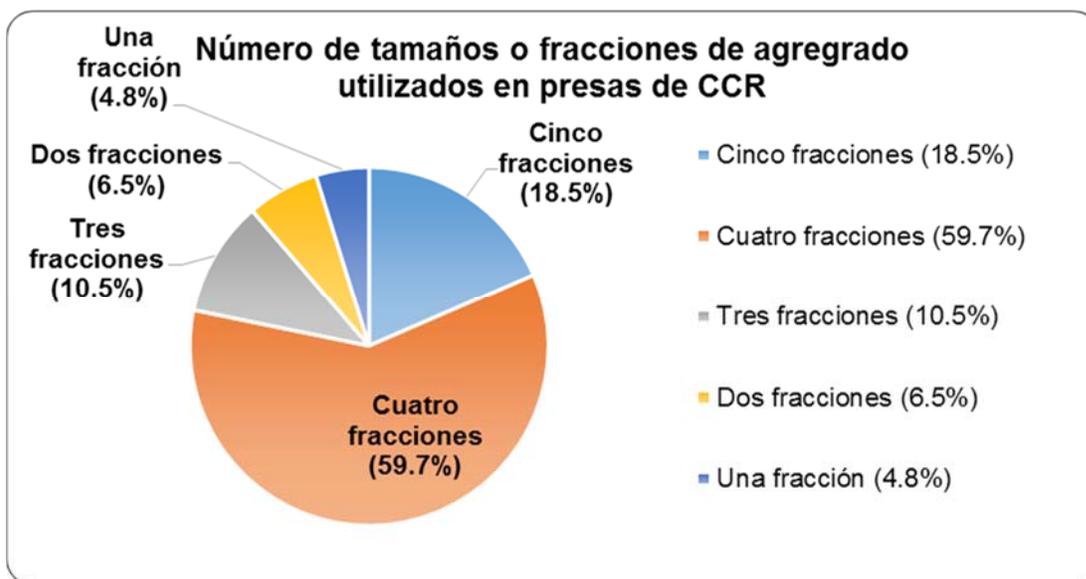


Figura 2-3 Número de tamaños o fracciones de agregado utilizados en presas de CCR.
(Tomado de ICOLD, 2003)

2.1.3. Aditivos

Los aditivos químicos más usados son, de acuerdo con la clasificación de ASTM C 494, los tipos A (reductor de agua), B (retardante) y D (reductor de agua y retardante). Estos aditivos al reducir la relación agua-cemento ayudan a obtener mayores resistencias y a mejorar la trabajabilidad.

El retardo de fraguado es importante para lograr conseguir juntas “calientes” entre las capas de CCR, y así, beneficiar la integración de capas subsecuentes sin la necesidad de colocar mezclas de liga entre ellas; la colocación de juntas incrementa el costo y retrasa los trabajos.

2.2. Selección de la dosificación de la mezcla.

Hay cuatro requisitos principales que debe cumplir un CCR para ser utilizado en estructuras de retención de agua y dos factores adicionales que se recomienda sean considerados. Los cuatro requisitos son: permeabilidad, densidad, resistencia y, la más importante, la capacidad de ser transportado, extendido y compactado sin segregación perjudicial. Los otros dos factores recomendables son: la durabilidad (ya que el CCR debe permanecer expuesto al ambiente sin estar completamente protegido mediante un revestimiento de concreto en los paramentos u otro tipo de protección) y las condiciones constructivas.

Como se ha mencionado, una de las preocupaciones principales en el diseño de CCR es la posibilidad de unión deficiente entre las capas. La consistencia de bajo revenimiento, particularmente en aquellos con agregados mayores a 38 mm y un contenido cementicio bajo, supone dificultad para unir el concreto fresco con concreto endurecido. Este problema se ha solucionado reduciendo el intervalo de tiempo entre la colocación de las capas, aplicando tratamientos de junta, o incrementando el contenido de pasta en la mezcla. Consecuentemente se debe garantizar la unión de las capas durante todas las etapas del proceso de diseño.

Otra posibilidad es que la sección transversal de la presa sea diseñada de tal manera que no se necesite adherencia en las juntas horizontales entre capas y se disponga algún tipo de membrana impermeable aguas arriba.

2.2.1. Permeabilidad.

Esta propiedad es la que ha creado mayor preocupación en los proyectistas de presas de CCR, ya que aunque se tenga una matriz de concreto con permeabilidad baja, son las juntas horizontales la causa principal de problemas en las presas. Se ha encontrado que la permeabilidad total *in situ* (incluyendo la de las juntas) oscila entre 10^{-4} y 10^{-13} m/s y se ha determinado que hay una estrecha relación entre la permeabilidad y el contenido cementicio, y en menor medida del uso de finos triturados.

Al parecer la consideración primordial que define la permeabilidad *in-situ* (incluida la de las juntas) de un CCR es el contenido cementicio total del concreto compactado con rodillo. Un CCR con un contenido cementicio más alto que crea una buena unión entre las capas producirá una permeabilidad adecuada. Sin embargo, la mayoría de las mezclas de CCR con contenidos cementicios menores no parecen producir la permeabilidad adecuada sin ningún tratamiento especial. Cuando se utiliza algún tipo de barrera de baja permeabilidad en el paramento aguas arriba para reducir la filtración en mezclas con

contenido cementicio bajo, la permeabilidad del CCR puede ser de poca importancia excepto en lo relativo a la durabilidad hielo-deshielo de las superficies expuestas.

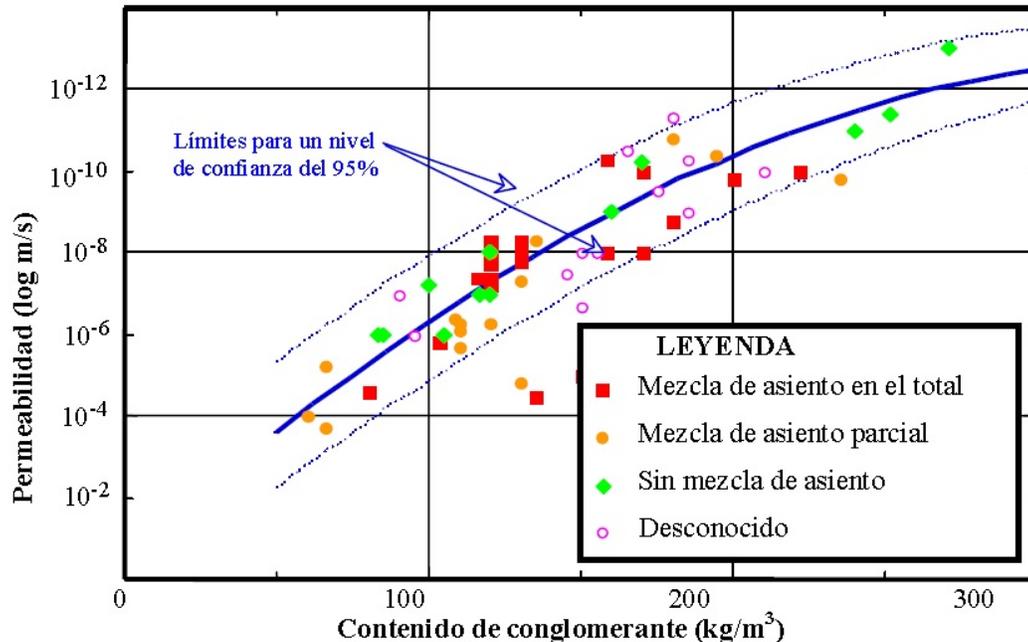


Figura 2-4 Permeabilidad vs. Contenido cementicio (tomado de ICOLD, 2003)

2.2.2. Densidad.

Aunque la densidad depende principalmente de la Densidad Relativa de los agregados utilizados en el concreto, el aire incluido conducirá a una pérdida de las propiedades. Puede obtenerse un amplio rango de densidades utilizando determinados materiales mezclados en diferentes proporciones. La *Figura 2-5* muestra una relación entre la relación pasta (material cementicio, agua y aire)-mortero (pasta más agregados finos) y la densidad en sitio. Se puede observar que hay una caída sustancial en la densidad por debajo de una relación de entre 0.35 a 0.40. Hay una estrecha correlación entre los resultados reales y los teóricos. Es posible añadir finos no plásticos al extremo final de la granulometría del agregado fino para reducir el índice de huecos; sin embargo, debe tenerse cuidado en este proceso ya que hay un punto en el que agregar más material fino puede incrementar el índice de vacíos.

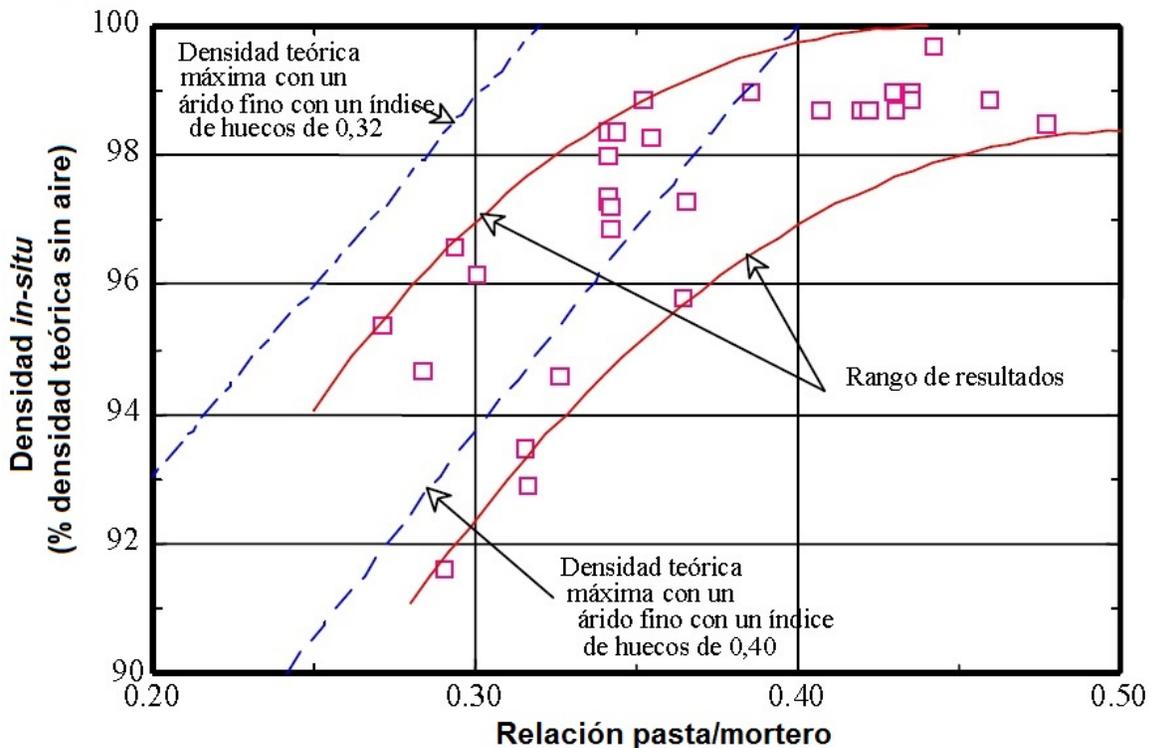


Figura 2-5 Densidad *in-situ* vs Relación pasta/mortero para mezclas de CCR
(Tomado de ICOLD, 2003)

2.2.3. Resistencia

La edad de diseño para mezclas de CCR se define a 91 o 180 días, dichos valores tienen en cuenta que la mezcla contiene adiciones minerales que desarrollan resistencia a largo plazo y que evidentemente habrá un incremento de la resistencia posterior a estas edades, el cual podrá ser tomado como un factor de seguridad adicional.

La resistencia de CCR depende de la calidad y granulometría de los agregados, de las proporciones del cemento, puzolana y agua, y del grado de compactación. Para la mayoría de las mezclas, la resistencia a compresión es inversamente proporcional a la relación agua/contenido cementicio w/c (el contenido cementicio es la suma del cemento y puzolana agregado a la mezcla), relación similar a la del concreto convencional.

La resistencia de diseño no se basa en la resistencia a compresión sino que depende más de las resistencias a tensión y cortante requeridas. La resistencia a compresión más bien es considerada como el método más conveniente para evaluar la calidad y uniformidad del concreto.

Las mezclas de CCR deben ser dosificadas para producir la resistencia a compresión de diseño más un margen basado en la variación esperada de la resistencia (de acuerdo con la *Tabla 2-3*).

Tabla 2-3 Coeficiente de Variación de varios tipos de ensayo en relación al nivel de control de calidad observado.
(Tomado ICOLD, 2003)

Evaluación del control de calidad a la edad de diseño)		Rangos de los Coeficientes de Variación (%)				
		Excelente	Bueno	Medio	Malo	Muy malo
Probetas						
Resistencia a compresión		< 10.0	10 a 15	15 a 20	20 a 25	> 25
Resistencia a tensión indirecta		< 12.5	12.5 a 17.5	17.5 a 22.5	22.5 a 27.5	> 27.5
Resistencia a tensión directa		< 15.0	15 a 20	20 a 25	25 a 30	> 30
Testigos						
Resistencia a compresión		< 15.0	15 a 20	20 a 25	25 a 30	> 30
Resistencia a tensión indirecta		< 17.5	17.5 a 25	25 a 32.5	32.5 a 40	> 40
Resistencia a tensión directa	sin junta	< 25.0	25 a 35	35 a 45	45 a 55	> 55
	con junta	< 30.0	30 a 40	40 a 50	50 a 60	> 60
Resistencia a cortante	sin junta	< 17.5	17.5 a 25	25 a 32.5	32.5 a 40	> 40
	con junta	< 20.0	20 a 27.5	27.5 a 35	35 a 42.5	> 42.5
	sin unión	< 22.5	22.5 a 30	30 a 37.5	37.5 a 45	> 45
Resistencia a cortante sobre testigo inclinado	sin junta	< 22.5	22.5 a 30	30 a 37.2	37.5 a 42.5	> 42.5
	con junta	< 25.0	25 a 32.5	32.5 a 40	40 a 47.5	> 47.5
Densidad						
Probetas		< 0.5	0.5 a 1.0	1.0 a 1.5	1.5 a 2.0	> 2.0
Ensayo VeBe		< 0.75	0.75 a 1.25	1.25 a 1.75	1.75 a 2.25	> 2.25
Densímetro nuclear		< 0.75	0.75 a 1.25	1.25 a 1.75	1.75 a 2.25	> 2.25
Testigos		< 1.0	1.0 a 1.5	1.5 a 2.0	2.0 a 2.5	> 2.5

2.2.4. Segregación

No importa cuán buenas sean las propiedades en laboratorio del CCR, si el material se segrega cuando se transporta, se extiende o se compacta, las propiedades *in situ* serán menores que aquellas obtenidas en las pruebas. Un objetivo prioritario en la dosificación de las mezclas de CCR es producir una mezcla cohesiva con la mínima tendencia posible a la segregación. Esto puede ser controlado hasta cierto punto mediante la granulometría del agregado y la adición de finos. La granulometría total y la forma del agregado deben ser seleccionadas cuidadosamente para maximizar la densidad del conjunto en estado suelto del agregado grueso para reducir así la posibilidad de segregación. Limitar el tamaño máximo de agregado también ayuda a reducir la tendencia a segregarse.

2.2.5. Trabajabilidad.

Se necesita una trabajabilidad adecuada para alcanzar la compactación o la consolidación de un CCR. La trabajabilidad es afectada sobre todo por la proporción de pasta de mezcla y por los finos de los agregados. La trabajabilidad de las mezclas se mide normalmente utilizando un aparato VeBe (ASTM, 1998) modificado o el ensayo VC (Nagayama, 1991).

Hay dos aparatos diferentes: el aparato estándar y el aparato mayor. El más utilizado es el dispositivo de ensayo VC con el contenedor estándar. En el ensayo VC estándar consiste en: extraer del concreto fresco mediante tamizado o cribado, los agregados mayores de 40 mm del RCD ("*Rolled Concrete Dam*" – método utilizado en Japón), el resto del concreto se coloca en el contenedor estándar. Una vez en el contenedor, el RCD se vibra.

El valor VC estándar se define como el tiempo en segundos en que la pasta cubre toda la superficie del RCD. El valor VC óptimo del RCD para compactación mediante rodillo vibratorio se ha determinado en 20 segundos.

El aparato mayor se utiliza para RCD con agregados hasta el tamaño máximo. A este valor VC se le llama el valor VC de gran tamaño. Se ha comprobado que un tiempo de vibrado de 60 segundos con el aparato mayor equivale a un tiempo de vibrado de 20 segundos con el aparato estándar.

Esta prueba se utiliza como un control similar al ensayo de revenimiento para el concreto convencional y adicionalmente da una medida de la densidad del concreto fresco. Las mezclas de CCR con el grado de trabajabilidad necesario para una fácil compactación y obtención de una densidad uniforme, para la unión con las capas previamente colocadas y para sustentar el equipo de compactación, generalmente tienen un tiempo VeBe de 10 a 30 segundos o un tiempo VC de 20 segundos. Sin embargo, las mezclas de CCR han sido dosificadas con un amplio rango de niveles de trabajabilidad alguno de los cuales contienen volúmenes de pasta tan bajos que la trabajabilidad no puede ser medida utilizando ningún ensayo diseñado para este efecto. Esto es particularmente cierto para aquellas mezclas dosificadas con un contenido cementicio muy bajo o para aquellas diseñadas para presas de suelo estabilizado con cemento (*hard-fill dams*) donde la trabajabilidad debe ser juzgada mediante observaciones durante la colocación y la compactación, junto con la medida de la densidad compactada y el contenido de humedad.

La demanda de agua para un nivel específico de trabajabilidad se verá influida por el tamaño, la forma, la textura y la granulometría de los agregados, y el volumen y la naturaleza de los materiales cementicios y de los finos.

2.2.6. Durabilidad.

La mezcla de CCR debe asegurar un cierto grado de durabilidad basándose en las condiciones de exposición, los materiales utilizados y el nivel que se espera de comportamiento (en diseño). El CCR debe estar libre de los efectos dañinos de la reactividad *álcali*-agregado mediante la adecuada evaluación y selección de los componentes.

2.2.7. Condiciones constructivas.

Durante el diseño de las dosificaciones se deben considerar los requisitos constructivos y la maquinaria disponible para la elaboración de mezclas *in situ*. Por ejemplo, si las temperaturas ambientes del lugar son particularmente altas y/o hay velocidades de viento elevadas, se debe tener especial cuidado en el diseño de la mezcla para que mantenga su trabajabilidad a lo largo del tiempo y para que tenga el máximo retardo posible de tal manera que cuando se coloquen las capas sucesivas se garantice la unión entre las juntas horizontales.

2.3. Métodos de dosificación de mezclas

Hay una serie de métodos que han sido utilizados para la selección de la dosificación en un CCR. La mayoría de ellos se basan en dos planteamientos generales: el enfoque de "concreto" en el cual se considera primordialmente la relación agua/cemento, y el enfoque de "suelos" en el cual la mezcla se diseña utilizando como aspecto central la relación humedad/densidad. Ambos enfoques tienen la intención de producir concretos de calidad adecuada para la compactación con rodillo y la construcción de grandes presas.

Las dosificaciones del CCR deben seguir el método utilizado en el concreto convencional, esto es, la identificación de la masa de cada ingrediente contenida en una la unidad de volumen compactada de la mezcla, basada en la condición de agregado saturado superficialmente seco (SSS). Se asumen agregados resistentes y limpios y que la resistencia del concreto completamente consolidado es inversamente proporcional a la relación agua–cemento (usando menos agua, con un contenido de cemento constante, se produce un concreto con mayor resistencia a la compresión y propiedades relacionadas).

El enfoque general está basado en el concepto de que existe suficiente pasta en la mezcla CCR que rellena todos los vacíos en el agregado, esto hace que la mezcla se compacte completamente produciéndose un concreto con revenimiento cero. Sin embargo, la mezcla CCR no debe contener un exceso de pasta que sea llevado a la superficie y se segregue con sólo unas pocas pasadas del rodillo vibratorio. En los últimos años ha habido una tendencia hacia el enfoque del "concreto", es decir, lograr CCR con contenidos cementicios mayores.

El procedimiento general de dosificación es el siguiente:

- i. Optimizar la granulometría de los agregados finos y grueso para producir los mínimos huecos en cada uno utilizando, si fuera necesario, finos adicionales en el agregado fino.
- ii. Escoger una relación pasta/mortero apropiada de tal manera que los vacíos en el agregado fino se llenen, o se sobrellenen ligeramente con la pasta (dicha relación se encuentra generalmente en el rango de 0.38 a 0.46, ver *Figura 2-5*).

- iii. Dosificar el cemento Portland, la adición mineral (si la hay), el agua y los aditivos (si los hay) para obtener la resistencia media requerida y para obtener la dosificación de la pasta.
- iv. Establecer el volumen de agregado grueso para obtener la trabajabilidad requerida, determinada mediante el tiempo VeBe.
- v. Comprobar que hay material cementicio suficiente (y una proporción de finos, si se utilizan) para conseguir la permeabilidad de diseño (*Figura 2-4*)
- vi. Comprobar que la relación agregado fino/agregado grueso es cercana al óptimo.
- vii. Comprobar que el calor de hidratación está dentro de los límites esperados.
- viii. Hacer cualquier ajuste que sea necesario y verificar el diseño.

- Dosificación de los agregados

Los agregados deben ser clasificados en varios tamaños o fracciones las cuales deben ser combinadas en diferentes proporciones hasta que se obtenga la densidad máxima en conjunto y así obtener la granulometría “ideal”. La densidad alta no es el único objetivo, también es probable que la segregación se reduzca incluso agregando materiales finos para reducir los vacíos.

- Dosificación de la pasta

El contenido de agua que se agregará a la mezcla está determinada por muchos factores y será únicamente válida para cada tipo estudio en particular. Se puede determinar el contenido de agua a utilizar si te tiene experiencia con los agregados específicos con los que se va a realizar el estudio (sitio) o, de lo contrario, se puede tomar como referencia inicial la mostrada en la *Tabla 2-4*.

*Tabla 2-4. Contenidos de agua aproximados para diferentes t.m.a.
(Tomado de ICOLD, 2003)*

Tamaño máximo de agregado (mm)	150	75	50	38	19
Contenido de agua libre (kg/m ³)	~85	~95	~100	~108	~117

- Determinación de la relación pasta/mortero

El mínimo de contenido cementicio, aire incluido y contenido de agua necesarios para alcanzar un volumen sólido, debe de rellenar todos los huecos del agregado fino y cubrir todas las partículas del agregado. El mínimo de volumen de pasta puede determinarse mediante las curvas de máxima densidad similares a las utilizadas para determinar el contenido óptimo de agua en suelos, esto es, para una volumen de pasta “inicial” se debe ir aumentando la cantidad de agregado fino en incrementos iguales al tiempo que se mide la densidad de las muestras después de compactadas. Con los datos obtenidos se traza la curva densidad-volumen de pasta y se determina la combinación que arroja la densidad máxima de la pasta.

- Dosificación del agregado grueso

El volumen mínimo de agregado que tiene una consistencia de revenimiento nulo puede ser establecido dosificando la fracción de mortero para obtener la resistencia necesaria aproximada y ajustando las proporciones de agregado grueso/mortero con las que se alcanza el revenimiento cero. Durante estos ajustes las proporciones del material cementicio, agua, agregado fino y del aire incluido deben mantenerse fijas

Una referencia, para aproximadamente definir el contenido de agregado grueso para CCR, se muestra en la *Tabla 2-5*:

*Tabla 2-5. Volumen absoluto de agregado grueso por unidad de volumen de concreto para diferentes t.m.a.
(Tomado de ICOLD, 2003)*

Tamaño máximo de agregado (mm)	150	75	50	38	19
Volúmen absoluto (m ³ /m ³)	0.63-0.64	0.57-0.61	0.54-0.58	0.52-0.56	0.46-0.52

2.4. Mezclas de prueba en laboratorio

Una vez que, con base en los puntos descritos anteriormente, se han definido los materiales a utilizar en el estudio (desde tipos de cementos, adiciones minerales, agregados disponibles y aprobados cercanos al sitio -derivado de un estudio de bancos previo), y se ha revisado la coherencia con las recomendaciones y referencias en el diseño de mezclas de CCR, se recomienda ampliamente que se dosifiquen e investiguen una serie de mezclas en el laboratorio para abarcar, en la mejor de las medidas, el rango potencial de características necesarias. Esta práctica permitirá modificaciones y/o ajustes posteriores (incluso durante la obra) a las dosificaciones de la mezcla sin tener necesariamente que repetir el programa de mezclas de ensayo.

Algunas de estas características son fácilmente determinadas mediante en examen visual de los ensayos de mezclas preparadas en el laboratorio. La dosificación del agregado en la mezcla, la “cohesividad” y la tendencia a la segregación pueden observarse mediante la manipulación del CCR en el laboratorio. La textura de la mezcla, ya sea áspera, no trabajable, arenosa, pastosa, suave, etc. se puede ver y sentir al tacto y pueden ser determinadas durante el estudio de diseño de mezclas. Estas características deben ser registradas para cada mezcla para permitir ajustes posteriores.

Las pruebas en estado fresco como temperatura, consistencia, peso unitario y contenido de aire, deben ser realizados en cada mezcla de ensayo. Además, se deben fabricar muestras para realizar ensayos de resistencias a compresión a varias edades, normalmente 7, 14, 28, 56, 90, 120, 180 días y 1 año para observar el desarrollo de la

resistencia conforme pasa el tiempo para cada mezcla. Estas muestras pueden ser también utilizadas para determinar otros parámetros como el módulo de elasticidad estático y el índice de Poisson a las edades seleccionadas. También se pueden fabricar muestras adicionales para la prueba de tensión indirecta (prueba Brasileña) y/o de resistencia a tensión directa a varias edades para establecer su relación con la resistencia a compresión y suministrar parámetros para usar en el análisis estructural.

Algunas muestras se pueden someter a curado acelerado en tanques para poder realizar los ensayos descritos y obtener resultados representativos de edades “lejanas” (un año) a corto plazo (~14 días).y con esta información tomar decisiones del camino a tomar durante el estudio.

De las mezclas “seleccionadas”, se pueden determinar otros parámetros importantes en el análisis estructural (no son parte del análisis que sostiene este trabajo de tesis) como: las propiedades térmicas, incluyendo el aumento adiabático de la temperatura, el coeficiente térmico de expansión, el calor específico y la difusividad, (se determinan normalmente para una o varias de las mezclas seleccionadas de CCR). Las muestras para ensayos especializados como son ensayos de fluencia, capacidad de deformación y resistencia a cortante, pueden ser desarrolladas únicamente en mezclas seleccionadas.

3. CAPÍTULO III. APRENDIZAJE MÁQUINA.

3.1. ¿Qué es el aprendizaje automático o aprendizaje máquina?

Sobre aprendizaje las definiciones de diccionario incluyen frases como “*para conseguir conocimiento sobre, comprensión de, o habilidad en ... mediante el estudio, instrucción, o experiencia*”, y “*modificación del comportamiento mediante la experiencia*”. Si estas descripciones se trasladan al entorno de las máquinas (hablando de procesos computacionales o desarrollados con computadoras) se puede decir, de manera muy amplia, que éstas aprenden cada vez que cambia su estructura, programa o datos internos con base en entradas o en respuesta a información externa (estímulos) de tal manera que el rendimiento esperado mejora.

El aprendizaje en una máquina generalmente se refiere a los cambios en los sistemas que realizan tareas relacionadas con la inteligencia artificial (IA). Estas tareas implican el reconocimiento, diagnóstico, planificación, control automatizado, predicción, etc. Los “cambios” pueden ser tanto mejoras en la ejecución de sistemas (tareas) o síntesis *ab initio* de nuevos sistemas. Para ser un poco más específicos, se muestra la arquitectura de un típico agente de IA en la *Figura 3-1*. En dicha Figura, el agente percibe y modela su entorno y calcula las acciones apropiadas, la mayoría de las veces mediante la anticipación de sus efectos. Los cambios realizados en cualquiera de los componentes mostrados en la Figura pueden contar como “aprendizaje”. Diferentes mecanismos de aprendizaje podrían emplearse dependiendo del subsistema que está siendo cambiado.

Para el lector interesado en el tema, descripciones más profundas pueden encontrarse e (Nilsson, 1998; Haykin, 2009).

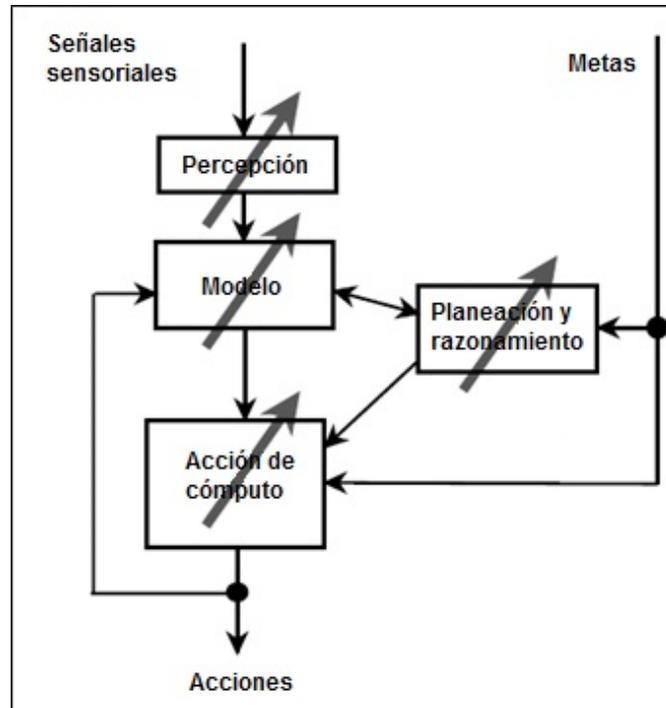


Figura 3-1 Un sistema de Inteligencia Artificial IA
(Tomado de Nilsson, 1998)

Razones importantes para la existencia del aprendizaje automático o máquina (AA) son:

- Algunas tareas no pueden ser definidas más que con ejemplos; es decir, podríamos ser capaces de especificar pares de entrada/salida, pero no una relación concisa entre las entradas y las salidas deseadas. El objetivo es que las máquinas sean capaces de ajustar su estructura interna para producir salidas correctas para un gran número de muestras de entrada y así limitar adecuadamente su función de entrada/salida para aproximarse a la relación implícita en los ejemplos.
- Es posible que ocultas entre grandes cantidades de datos existan relaciones y correlaciones importantes. Los métodos del aprendizaje automático a menudo pueden utilizarse para obtener estas relaciones (minería de datos).
- Los diseñadores humanos a menudo producen máquinas que no funcionan tan bien como se desea en el entorno donde son utilizadas. De hecho, es posible que ciertas características del entorno de trabajo no se conozcan por completo al momento del diseño. Los métodos del aprendizaje automático también se pueden utilizar para la mejora de la operación de máquinas que ya existen.
- La cantidad de conocimiento disponible, sobre ciertas tareas, puede ser demasiado grande para la codificación explícita hecha por personas. Las máquinas que aprenden este conocimiento gradualmente podrían ser capaces de capturar más de lo que las personas querrían o serían capaces de "anotar".

- Los ambientes cambian con el tiempo. Las máquinas que pueden adaptarse a un entorno cambiante reducirían la necesidad de rediseño constante y esto hace eficientes a los procesos.
- Nuevo conocimiento sobre todos los temas que atiende la humanidad está siendo descubierto o creado constantemente por ejemplo: vocabulario, corrientes o patrones de acontecimientos, fenómenos naturales, etc. El continuo rediseño de sistemas de IA para ajustarse a los nuevos conocimientos no es práctico, pero los métodos del aprendizaje automático podría ser capaz de realizar un seguimiento de gran parte de ellos.

Componentes del Aprendizaje Automático

El AA puede ser visto como la convergencia de varias fuentes científicas en la que cada disciplina aporta sus propios métodos y vocabulario para ser asimilados y unificados en un método. Algunas de las disciplinas que participan son:

- Estadística.
- Modelos cerebrales (neurociencias).
- Teoría del control adaptativo.
- Modelos Psicológicos.
- Inteligencia Artificial.
- Modelos Evolutivos.

Variantes del AA

Un objetivo del AA es aprender una estructura computacional de algún tipo, cualquiera, para lo cual se consideran una variedad de diferentes estructuras de cálculo:

- Funciones
- Programas de lógica y conjuntos de reglas
- Máquinas de estados finitos
- Gramática
- Sistemas de resolución de problemas

3.1.1. Aprendizaje de funciones de entrada-salida

Tipos de aprendizaje

Hay dos escenarios importantes en los que queremos aprender una función. En uno, llamado *aprendizaje supervisado*, conocemos los valores de f para las m muestras en el *conjunto de entrenamiento* $[x]$. Suponemos que si podemos encontrar una hipótesis, h , que coincida estrechamente con f para los miembros de $[x]$, entonces esta hipótesis será una buena suposición para f , especialmente si $[x]$ es grande.

Un ajuste de una función mediante una curva es un ejemplo sencillo de aprendizaje supervisado, como la que se muestra en la *Figura 3-2*. La función parabólica, h , mostrada en tal figura, es una hipótesis acerca de una función f que produjo las cuatro muestras mostradas.

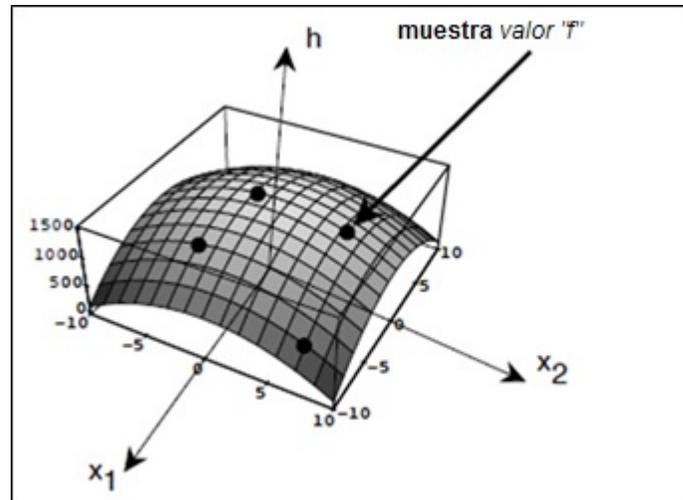


Figura 3-2. Ejemplo de un ajuste mediante una función parabólica.
(Tomado de Nilsson, 1998)

En el otro escenario, denominado *aprendizaje no supervisado*, simplemente tenemos un conjunto de entrenamiento de vectores sin valores de función para ellos. El problema en este caso, por lo general, es dividir el conjunto de entrenamiento en subconjuntos, $[x]_1, \dots, [x]_R$, de alguna manera apropiada.

También existen métodos que son intermedios entre el aprendizaje supervisado y no supervisado al tratar de encontrar una nueva función, h , o modificar una existente. Un caso especialmente interesante es el de cambiar una función existente en una equivalente que es computacionalmente más eficiente. Este tipo de aprendizaje es a veces llamado *aprendizaje de aceleración*.

Un ejemplo muy simple de aprendizaje de aceleración implica procesos de deducción. A partir de las fórmulas $A \supset B$ y $B \supset C$, podemos deducir C si se conoce A . A partir de este proceso deductivo podemos crear la fórmula $A \supset C$ (una nueva fórmula pero que no sanciona cualquier conclusión que podría derivarse de las fórmulas que teníamos anteriormente). Pero con esta nueva fórmula podemos derivar C más rápidamente, dado A , lo que podríamos haber hecho antes. Podemos contrastar el aprendizaje acelerado con métodos que crean realmente nuevas funciones (unos que podrían dar resultados diferentes después del aprendizaje de lo que hacían antes). Decimos que los últimos métodos implican *aprendizaje inductivo*. A diferencia de la deducción, no hay inducciones correctas sólo más útiles.

Vectores de entrada

El vector de entrada es llamado por una amplia variedad de nombres. Algunos de éstos son: *vector de entrada*, *vector patrón/patrones*, *vector de características*, *muestras*, *ejemplos*, e *instancias*. Los componentes, x_i , del vector de entrada se llaman igualmente, de muy diversas maneras: *características*, *atributos*, *variables de entrada*, y *componentes*.

Los valores de los componentes pueden ser de tres tipos principales: números con resultados reales (valuados), números discretos valuados o valores categóricos.

Salidas

La salida puede ser un número real, caso en el que el proceso que realiza la función, h , se llama estimador de la función y a la salida se le denomina un valor de salida o estimación. Alternativamente, la salida puede ser un valor categórico, en cuyo caso el proceso que incorpora h se llama indistintamente *clasificador*, *reconocedor*, o *categorizador*, y a la salida se le llama *etiqueta*, *clase*, *categoría* o *decisión*. Salidas valoradas por vectores también son posibles con componentes de números reales o valores categóricos.

Un caso especial importante es el de valores booleanos de salida. En este caso, un patrón de entrenamiento que tiene un valor 1 se denomina una *instancia positiva* y una muestra de entrenamiento que tiene un valor 0 se denomina una *instancia negativa*. Cuando la entrada es también booleana el clasificador implementa una función booleana. Estudiar el caso booleano con cierto detalle, permite hacer observaciones generales importantes en un entorno simplificado. El aprendizaje de una función lógica a veces se llama el *aprendizaje de conceptos*, y la función se llama un *concepto*.

Regímenes de entrenamiento

Hay varias formas en las que el conjunto de entrenamiento, $[x]$, se puede utilizar para producir una función hipotética. En el método *discontinuo*, todo el conjunto de entrenamiento está disponible y se utiliza para calcular la función, h . Una variación de este método utiliza todo el conjunto de entrenamiento para modificar una hipótesis inicial iterativamente hasta que se obtiene una hipótesis aceptable.

Por el contrario, en el método *gradual* se selecciona un miembro a la vez del conjunto de entrenamiento y usamos este caso para modificar una hipótesis inicial. Luego, otro miembro del conjunto de entrenamiento se selecciona y así sucesivamente. El método de selección puede ser aleatorio (con reemplazo) o puede desplazarse por el conjunto de entrenamiento de forma iterativa. Si todo el conjunto de entrenamiento dispone de un miembro a la vez, entonces también podríamos utilizar un método gradual, es decir seleccionar y utilizar miembros del conjunto de entrenamiento a medida que “llegan” o se presentan.

Por otra parte, en cualquier etapa todos los miembros del conjunto de la formación, hasta ahora disponibles, podrían utilizarse en un proceso discontinuo. Utilizar los miembros del conjunto de entrenamiento a medida que estén disponibles se llama método en línea. Los métodos en línea podrían utilizarse, por ejemplo, cuando la siguiente instancia de entrenamiento es una función de la hipótesis actual y la instancia anterior.

Contaminación/Ruido

A veces los vectores en el conjunto de entrenamiento están contaminados por ruido. Hay dos tipos de ruido. El *ruido en clases* altera aleatoriamente el valor de la función; *ruido en atributos* altera aleatoriamente los valores de las componentes del vector de entrada. En cualquier caso no es apropiado insistir que la función hipotética concuerda precisamente con los valores de las muestras en el conjunto de entrenamiento.

Evaluación del Desempeño

A pesar de que no hay una respuesta correcta en el aprendizaje inductivo es importante disponer de métodos para evaluar el resultado del aprendizaje. De manera resumida, en el aprendizaje supervisado la función inducida generalmente se evalúa en un conjunto separado de las entradas y valores de la función llamado *conjunto de prueba*. Una función hipotética es llamada *generalizada* cuando predice bien las respuestas conocidas del conjunto de pruebas. Tanto el error cuadrático medio como el número total de errores son medidas comunes en este proceso.

3.1.2. El aprendizaje y el sesgo.

Ciertamente hay un número incontable de funciones diferentes que tienen valores que coinciden con los cuatro ejemplos (puntos) mostrados en la *Figura 3-2*. Con esta selección de patrones por ajustar, la selección que naturalmente surge para limitar *a priori* el conjunto de hipótesis (en este caso funciones cuadráticas) y que luego podría ser ampliamente probado (mostrar que una función, la “elegida”, pasa a través de los cuatro puntos de muestreo) se llama sesgo. Es importante reconocerlo y ubicarlo apropiadamente dentro del marco de conocimiento adquirido, pues el aprendizaje sin sesgo es imposible, incluso muchas veces puede resultar útil.

Veamos un ejemplo específico de cómo el sesgo ayuda al aprendizaje. Una función booleana puede ser representado por un hipercubo donde cada uno de los vértices representa un patrón de entrada diferente. En la *Figura 3-3* se muestra una versión de 3 dimensiones en el que aparece un *conjunto de entrenamiento* de seis patrones (muestras), marcados los que tienen un valor de 1 por un punto cuadrado y los que tienen un valor de 0 por uno circular. Si el conjunto de hipótesis se compone de sólo las funciones linealmente separables (es decir, aquellas para las que los casos positivos y negativos pueden ser separados por una superficie lineal) entonces no es sólo una la función que en esta hipótesis es consistente con el *conjunto de entrenamiento*. Así que, en este caso, a pesar de que el conjunto de entrenamiento no contiene todos los patrones posibles, ya podemos precisar lo que la función debe ser (dado el sesgo).

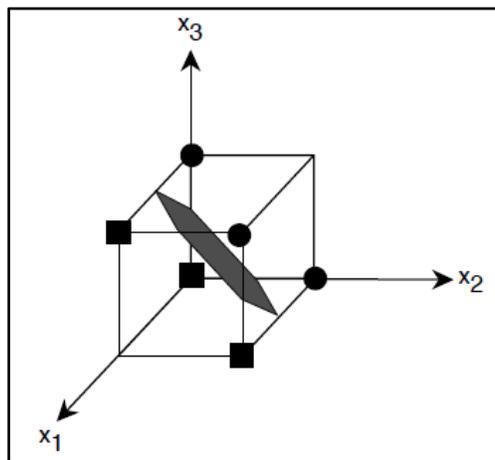


Figura 3-3. Un conjunto de entrenamiento o base de datos que determina completamente una función linealmente separable (Tomado de Nilsson, 1998)

Investigadores del “aprendizaje automático” han identificado dos variedades principales de sesgo, el *absoluto* y el *preferencial*. En el *sesgo absoluto* (también llamado sesgo con espacio de hipótesis restringido) se restringe H a un subconjunto definido de funciones. En nuestro ejemplo de la *Figura 3-3* la restricción fue a las funciones Booleas linealmente separables. En el *sesgo preferencial* se selecciona el mínimo número de hipótesis sobre todas las posibles de acuerdo con algún esquema de ordenamiento. Por ejemplo, si tuviéramos alguna manera de medir la complejidad de una hipótesis, podríamos seleccionar la que fue más simple entre las que ejecutan de manera satisfactoria en el conjunto de entrenamiento.

A continuación se describe brevemente las características del algoritmo del aprendizaje máquina que es usado en el estudio específico motivo de este trabajo de tesis. Textos más profundos sobre la matemática y las pruebas de lo anteriormente descrito se hallan en (Hertz, Krogh, y Palmer, 1991; Weiss y Kulikowski, 1991; Natarajan, 1991; Fu, 1994; Langley, 1996; Nilsson, 1998; Witten, y Frank, 2005; Haykin, 2009; Maimon, Rokach, 2010; Liao, et al, 2012). Existen además conferencias y publicaciones relacionadas como son:

- *Conferencia Anual sobre Avances en Sistemas de Procesamiento de Información Neuronal (NIPS por sus siglas en inglés) "*
- *Seminario Anual sobre la Teoría del Aprendizaje Computacional*
- *Conferencia Internacional Anual sobre Aprendizaje Automático (Machine Learning)"*
- *Conferencia Internacional Anual sobre Algoritmos Genéticos*
- *La revista o publicación sobre Aprendizaje Automático (The journal Machine Learning)*

3.2. Árboles de decisión

Un árbol de decisión es un árbol cuyos nodos internos (o ramas) son pruebas (para los patrones de entrada) y sus hojas son categorías (de los patrones). En la *Figura 3-4* se muestra un ejemplo. Un árbol de decisión asigna un número de clase (o salida) a un patrón de entrada al filtrarlo a través de las ramas del árbol. Cada prueba tiene resultados exhaustivos y mutuamente excluyentes. Por ejemplo, probar T_2 en el árbol de la *Figura 3-4* tiene tres resultados; a la izquierda se asigna el patrón de entrada a la clase 3, el de en medio envía el patrón de entrada hacia abajo para poner a prueba T_4 , y a la derecha se asigna el patrón a la clase 1. Seguimos la convención usual de representar los nodos hoja por el número de clase. La discusión (presentación) sobre los árboles de decisión no se limita a la implementación de funciones Booleas (las cuales por lo general son útiles para funciones valuadas categóricamente).

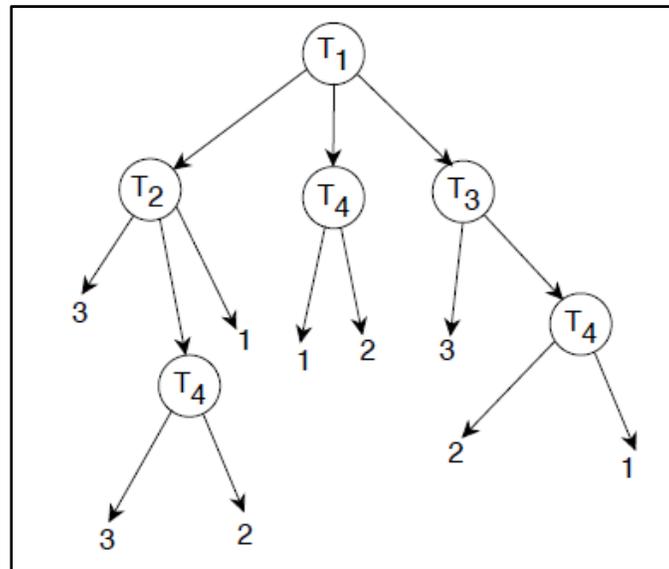


Figura 3-4. Esquema de árbol de decisión.
(Tomado de Nilsson, 1998)

Hay varias dimensiones en las que los árboles de decisión pueden diferir:

- Las pruebas podrían ser *multivariadas* (realizando pruebas de varias características de una entrada a la vez) o *univariado* (probando una sola de las características).
- Las pruebas pueden tener dos o más resultados (si todas las pruebas tienen dos resultados posibles se habla de un árbol de decisión binaria).
- Las características o atributos podrían ser categóricos o numéricos (los de valor binario pueden ser considerados como cualquiera de los dos).
- Podríamos tener dos o más clases (si sólo hay dos clases y entradas binarias el árbol implementa una función Boleana y se llama un árbol de decisión de Boleana).

Es sencillo representar la función implementada por un árbol de decisión Boleana univariado en forma de DNF (Forma Normal Disyuntiva, siglas en inglés *Disjunctive Normal Form*). La forma DNF implementada por un árbol de este tipo puede obtenerse mediante:

i) trazando hacia abajo de cada camino que conduce a un nodo de punta correspondiente a un valor de salida de 1, ii) formando el conjunto de las pruebas a lo largo de este camino, y luego iii) tomando la disyunción de estas conjunciones. En la *Figura 3-5* se muestra un ejemplo. En la elaboración de los árboles de decisión univariados, cada nodo “no-hoja” se representa por un único atributo. Si el atributo tiene el valor 0 en el patrón de entrada, se toma la rama por la izquierda; si tiene valor 1, se debe dirigir por la rama de la derecha.

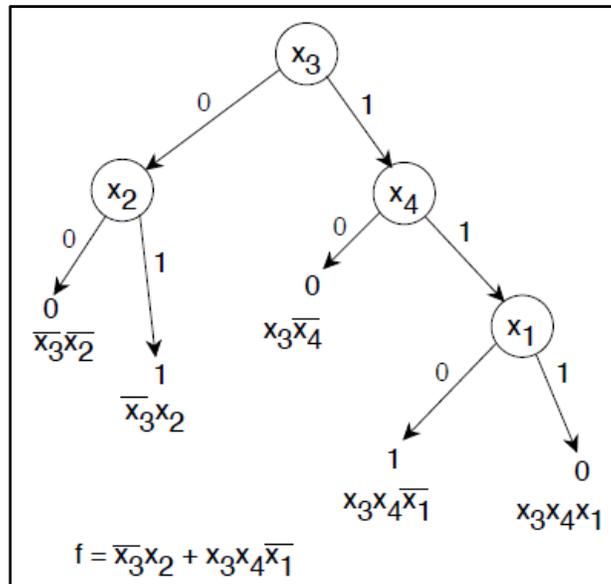


Figura 3-5. Un árbol de decisión implementando una función DNF
(Tomado de Nilsson, 1998)

La clase k-DL de funciones Booleanas se puede implementar mediante un árbol de decisión multivariante que tiene la forma (desequilibrada) mostrada en la Figura 3-6. Cada prueba, c_i , es un término de tamaño k o menor. Toda v_i tiene valores de 0 o 1.

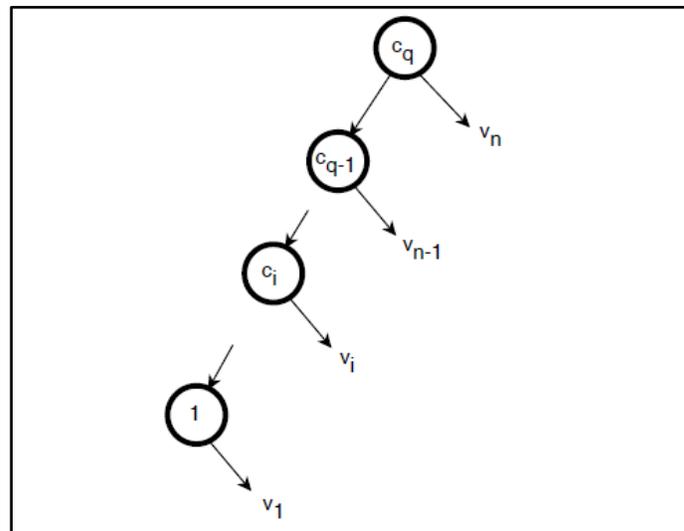


Figura 3-6. Un árbol de decisión implementando una lista de decisión
(Tomado de Nilsson, 1998)

3.3. Aprendizaje supervisado de árboles de decisión univariados

Se han propuesto varios sistemas de árboles de decisión de aprendizaje. Entre ellos destacan los ID3 y su nueva versión, C4.5 (Quinlan, 1986, Quinlan, 1993), y CART (Breiman, et al., 1984).

Selección del tipo de prueba

Como ha sido descrito anteriormente, se parte de n características o atributos. Si los atributos son binarios, las pruebas son simplemente si el valor del atributo es 0 o 1. Si los atributos son categóricos, pero no binarios, las pruebas podrían formarse dividiendo los valores de los atributos en subconjuntos mutuamente excluyentes y exhaustivos. Un árbol de decisión con estas pruebas se muestra en la *Figura 3-7*. Si los atributos son numéricos, las pruebas podrían implicar “pruebas de intervalo”, por ejemplo $7 \leq x_i \leq 13.2$.

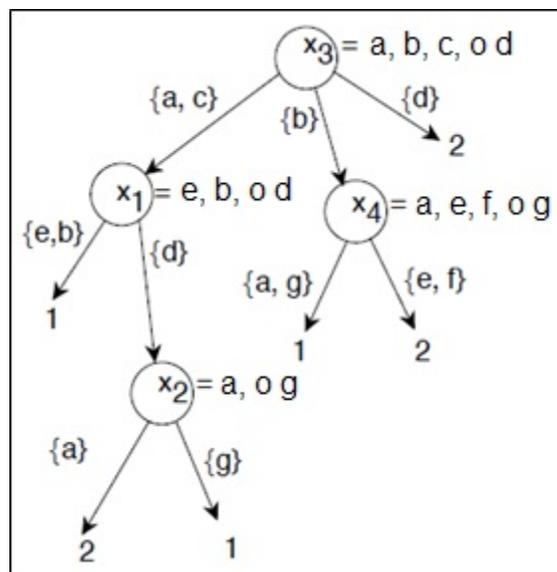


Figura 3-7. Árbol de decisión con atributos categóricos
(Tomado de Nilsson, 1998)

Uso de la reducción de la incertidumbre para seleccionar pruebas.

El principal problema de aprendizaje en árboles de decisión para el caso de atributos binarios es seleccionar el orden de las pruebas. Para los atributos numéricos y categóricos, también tenemos que decidir cuáles deben ser las pruebas (además de seleccionar el orden). Varias técnicas han sido probadas siendo una de las más populares seleccionar en cada etapa la prueba que reduce al máximo las entropías similares medidas.

Para dar una idea del método se ejemplifica un caso simple de pruebas con resultados binarios para desarrollar “manualmente”. Los casos más complejos deben analizarse con la ayuda de herramientas computacionales.

La entropía o incertidumbre sobre la clase de un patrón, que pertenece a algún conjunto $[x]$ de patrones se define como:

$$H[x] = - \sum_i p(i|[x]) \log_2 p(i|[x]) \quad \text{Ecuación 3-1}$$

Donde $p(i|[x])$ es la probabilidad de que un patrón seleccionado al azar de $[x]$ pertenece a la clase i y la sumatoria es sobre todas las clases. El objetivo es seleccionar las pruebas en cada nodo de tal manera que mientras se desciende por el árbol, la incertidumbre acerca de la clase de un patrón se vuelve cada vez menor.

El término $p(i|[x])$ puede ser estimado de las estadísticas de la muestra. Suponiendo que $\hat{p}(i|[x])$ es el número de patrones en $[x]$ que pertenece a la clase i dividido por el número total de patrones en $[x]$, entonces podemos estimar la incertidumbre como:

$$\hat{H}[x] = - \sum_i \hat{p}(i|[x]) \log_2 \hat{p}(i|[x]) \quad \text{Ecuación 3-2}$$

Si realizamos una prueba, T , que tiene resultados posibles k en los patrones en $[x]$, crearemos k subconjuntos, $[x]_1, [x]_2, \dots, [x]_k$. Supongamos que n_i de los patrones en $[x]$ están en $[x]_i$ para $i = 1, \dots, k$. (algunos pueden ser 0). Si supiéramos que T aplica a un patrón en $[x]$ dando lugar a la obtención de resultados j (es decir, si se sabe que el patrón estaba en $[x]_j$), la incertidumbre acerca de su clase será:

$$H[x_j] = - \sum_i p(i|[x_j]) \log_2 p(i|[x_j]) \quad \text{Ecuación 3-3}$$

Y la reducción en la incertidumbre será:

$$H[x] - H[x_j] \quad \text{Ecuación 3-4}$$

Por supuesto, no se puede decir que la prueba T garantiza siempre producir esa cantidad de reducción de la incertidumbre porque no se sabe si el resultado de la prueba será el resultado del j -ésimo; pero podemos estimar la incertidumbre media durante todo el $[x]_j$, por:

$$E[H_T([x])] = \sum_j p([x]_j) H([x]_j) \quad \text{Ecuación 3-5}$$

Donde por $H_T([x])$ nos referimos a la incertidumbre media después de realizar la prueba T de los patrones en $[x]$, $p([x]_j)$ es la probabilidad de que la prueba tiene un resultado j , y la suma se toma de 1 a k . Una vez más, no sabemos la probabilidad $p([x]_j)$, pero podemos usar valores de la muestra. La estimación $\hat{p}([x]_j)$ de $p([x]_j)$ es sólo el número de esos patrones en $[x]$ que tienen resultado j dividido por el número total de patrones en $[x]$. La reducción media de la incertidumbre lograda mediante la prueba T (aplicado a los patrones en $[x]$) es entonces:

$$R_T([x]) = H([x]) - E[H_T([x])]$$

Ecuación 3-6

Una importante familia de algoritmos de aprendizaje de árboles de decisión selecciona para la raíz del árbol la estructura que pruebe la máxima reducción de la incertidumbre. Aplica este criterio de forma recursiva hasta que se cumpla alguna condición de terminación (que discutiremos en detalle más adelante). Los cálculos de incertidumbre son particularmente simples cuando las pruebas tienen resultados binarios y cuando los atributos tienen valores binarios. A continuación se presenta un ejemplo sencillo para ilustrar cómo funciona el mecanismo de selección de la prueba descrita en este caso.

Supongamos que queremos usar el método de reducción de la incertidumbre para construir un árbol de decisión para clasificar los siguientes patrones:

Patrón	Clase
(0, 0, 0)	0
(0, 0, 1)	0
(0, 1, 0)	0
(0, 1, 1)	0
(1, 0, 0)	0
(1, 0, 1)	1
(1, 1, 0)	0
(1, 1, 1)	1

¿Qué prueba única, x_1 , x_2 o x_3 , se debe realizar en primer lugar? La ilustración en la Figura 3-8 da intuición geométrica sobre el problema.

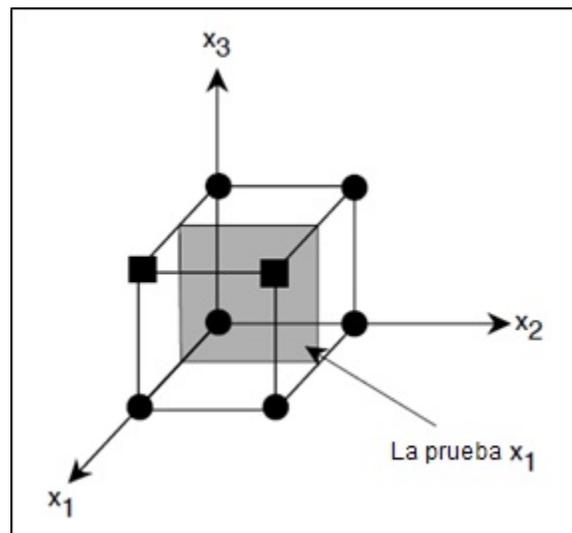


Figura 3-8. Ocho patrones a ser clasificados mediante un árbol de decisión.
(Tomado de Nilsson, 1998)

La incertidumbre inicial para el conjunto, que contiene a los ocho puntos es:

$$H([x]) = -(6/8)\log_2(6/8) - (2/8)\log_2(2/8) = 0.81$$

A continuación, se calcula la reducción de la incertidumbre. Si realizamos x_1 primero, la rama de la izquierda sólo tiene patrones que pertenecen a la clase 0 (los llamamos el conjunto $([x]_l)$), y la rama derecha $([x]_r)$ tiene dos patrones en cada clase. Por lo tanto, la incertidumbre de la rama izquierda es:

$$H_{x_1}([x]_l) = -(4/4)\log_2(4/4) - (0/4)\log_2(0/4) = 0$$

Y la incertidumbre de la rama derecha es:

$$H_{x_1}([x]_r) = -(2/4)\log_2(2/4) - (2/4)\log_2(2/4) = 1$$

La mitad de los patrones "van a la izquierda" y la otra mitad "van a la derecha" en la prueba x_1 . Por lo tanto, la incertidumbre media después de realizar la prueba x_1 es:

$$1/2 H_{x_1}([x]_l) + 1/2 H_{x_1}([x]_r) = 0.5$$

Así la reducción de la incertidumbre sobre $[x]$ alcanzada por x_1 es:

$$R_{x_1}([x]) = 0.81 - 0.5 = 0.31$$

Al hacer los mismos cálculos, vemos que la prueba x_3 alcanza exactamente la misma reducción de la incertidumbre, pero x_2 no logra reducción alguna. Por lo tanto, nuestro algoritmo para la selección de la primera prueba podría ser bien x_1 o x_3 . Supongamos que elegimos realizar primero x_1 , el proceso de reducción de la incertidumbre seleccionaría x_3 como la siguiente prueba. Así, el árbol de decisión que este procedimiento desarrolla implementa la función Boleana: $f = x_1x_3$.

Atributos no binarios

Aún si los atributos no son binarios se puede utilizar la técnica de reducción de la incertidumbre para seleccionar pruebas. La diferencia es que, además de seleccionar un atributo, debemos seleccionar una prueba en ese atributo. Supongamos por ejemplo que el valor de un atributo es un número real y que la prueba a realizar es establecer un umbral y ponerlo a prueba para ver si el número es mayor o menor que el umbral. En principio, dado un conjunto de patrones marcados, podemos medir la reducción de la incertidumbre para cada prueba que se consigue por cada umbral posible (sólo hay un número finito de umbrales que dan diferentes resultados de la prueba si sólo hay un número finito de patrones de entrenamiento). Del mismo modo, si un atributo es categórico (con un número finito de categorías), sólo hay un número finito de subconjuntos mutuamente excluyentes y exhaustivos en la que los valores del atributo se pueden dividir. Podemos calcular la reducción de la incertidumbre para cada división.

Redes Equivalente a árboles de decisión

Dado que los árboles de decisión booleanos univariados son implementaciones de las funciones de DNF, también son equivalentes a las redes neuronales de alimentación al frente de dos capas (Haykin, 2009). En la *Figura 3-9* se muestra un ejemplo. El árbol de

decisión de la izquierda se desarrolló implementando la misma función que la red a la derecha de la Figura. Por supuesto, cuando se implementa como una red, todas las características se evalúan en paralelo para cualquier patrón de entrada, mientras que cuando se implementa como un árbol de decisión sólo aquellas características en la rama se mueven hacia abajo por la necesidad de evaluar al patrón de entrada.

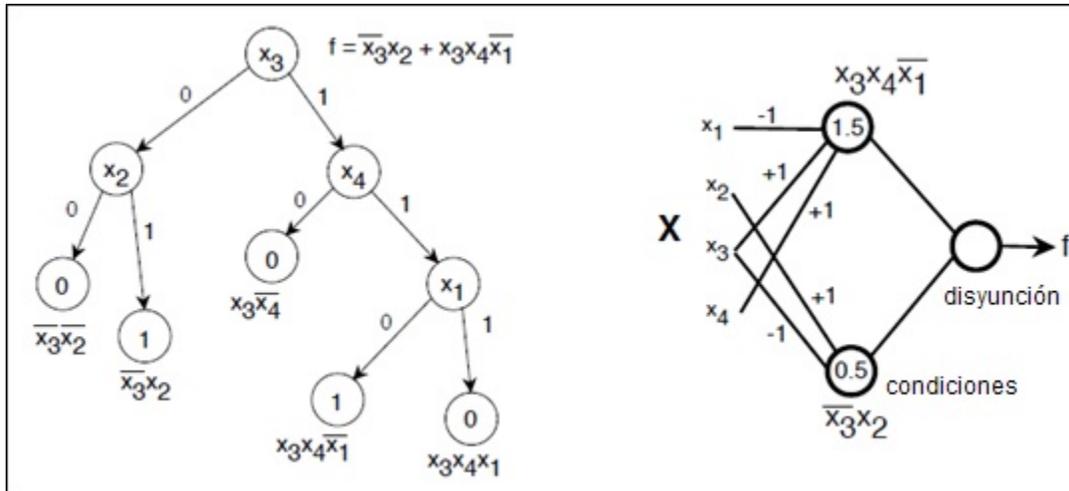


Figura 3-9. Árbol de decisión univariado y su red equivalente
(Tomado de Nilsson, 1998)

Los árboles de decisión multivariados con funciones linealmente separables en cada nodo también pueden ser implementados con redes de alimentación al frente –en este caso de tres capas-. En la Figura 3-10 se muestra un ejemplo en la que las funciones linealmente separables, cada una que utiliza un TLU (*Threshold Logic Unit* o unidad lógica de umbral), se indican con L_1 , L_2 , L_3 , y L_4 . Una vez más, la última capa tiene pesos fijos, pero los pesos en las dos primeras capas deben ser entrenados. Diferentes enfoques de los procedimientos de formación han sido discutidos por (Brent, 1990), por (John, 1995), y para un caso especial por (Marchand y Golea, 1993).

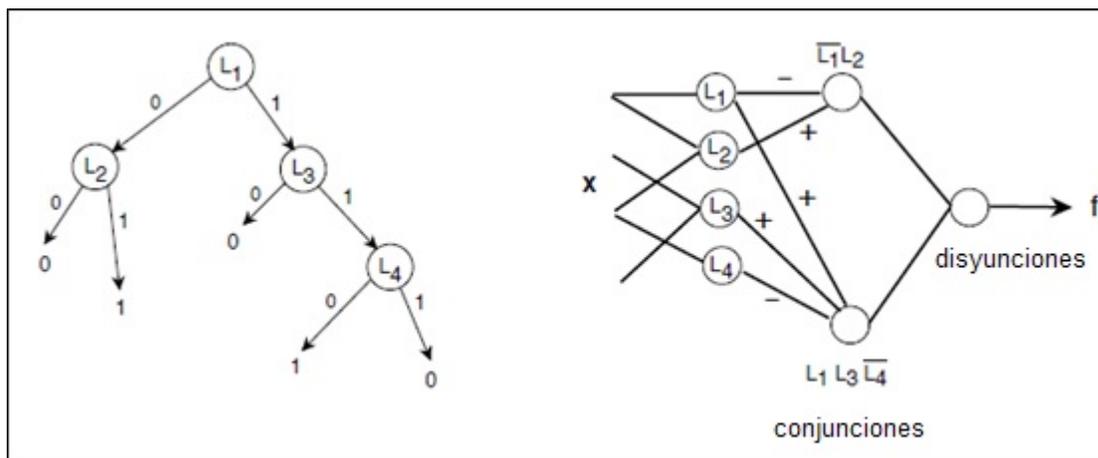


Figura 3-10. Árbol de decisión multivariado y su red equivalente.
(Tomado de Nilsson, 1998)

Memorización y evaluación

En el aprendizaje supervisado se debe elegir una función para ajustar el conjunto de entrenamiento de entre un conjunto de hipótesis. Ya hemos demostrado que la generalización es imposible sin sesgo. Cuando conocemos *a priori* que la función que estamos tratando de adivinar pertenece a un pequeño subconjunto de todas las funciones posibles, entonces, incluso con un conjunto incompleto de muestras de entrenamiento, es posible reducir suficientemente el subconjunto de funciones de acuerdo con el conjunto de entrenamiento para hacer conjeturas útiles sobre el valor de la función de las entradas que no están en el conjunto de entrenamiento. Cuanto mayor sea el conjunto de entrenamiento lo más probable es que incluso una función consistente seleccionada al azar tendrá salidas apropiadas para los patrones que no se muestran en el entrenamiento (o proceso de aprendizaje).

Sin embargo, incluso con sesgo, si el conjunto de entrenamiento no es lo suficientemente grande en comparación con el tamaño del espacio de hipótesis, habrá demasiadas funciones consistentes para que hagamos conjeturas útiles y la generalización de rendimiento será pobre. Cuando hay demasiadas hipótesis que sean coherentes con el *conjunto de entrenamiento* se dice que se está sobre-ajustando a los datos de entrenamiento. Sobre-ajustar, o memorizar, es un problema que hay que abordar en todos los métodos de aprendizaje. Se han propuesto varias técnicas para evitar la memorización, las cuales se describen a continuación.

Métodos de Validación

La forma más sencilla de estimar la aptitud de una función hipótesis (en el caso de esta investigación, un árbol de decisión) se realizan ensayos en unidades de prueba. Pero, si se comparan varios sistemas de aprendizaje, para seleccionar al más eficiente, ante un conjunto de pruebas, entonces tal comparación significa “entrenar con los datos de prueba”.

A continuación describimos algunas técnicas de validación que tratan de evitar los problemas de memorización.

Validación cruzada

En la validación cruzada se divide al conjunto de entrenamiento $[x]$ en K subconjuntos mutuamente excluyentes y exhaustivos de igual tamaño: $[x]_1, \dots, [x]_K$. Para cada subconjunto, $[x]_i$, se entrena en la unión de todos los otros subconjuntos y empíricamente se determina la tasa de error, ε_i , en $[x]_i$. La tasa de error es el número de errores de clasificación realizados en $[x]_i$ dividido por el número de patrones en $[x]_i$. Una estimación de la tasa de error que se puede esperar en los nuevos patrones de un clasificador entrenado en todos los patrones en $[x]$ es entonces el promedio de la ε_i .

Validación “de una salida”

La validación “de una salida” es la misma que la validación cruzada aplicada al caso especial en el que K es igual al número de patrones en $[x]$, y cada $[x]_i$ consta de un único patrón. Cuando se prueba en cada $[x]_i$, simplemente se observa si se cometió un error. Se cuenta el número total de errores y se divide entre K para obtener el porcentaje de error estimado. Este tipo de validación es, por supuesto, más difícil computacionalmente, pero útil cuando se necesita una estimación más precisa de la tasa de error para un clasificador.

Evitar memorización en árboles de decisión

Cerca de la parte más alta de un árbol de decisión puede haber sólo unos pocos patrones por nodo. Para estos nodos, estamos seleccionando una prueba basada en una muestra muy pequeña, y por lo tanto es probable que memorice. Este problema puede ser tratado mediante finalización del desarrollo de la prueba antes que todos los patrones están perfectamente divididos en sus distintas categorías. Este procedimiento dará lugar a algunos errores, pero a menudo se acepta un pequeño número de errores en los resultados en un conjunto de pruebas.

Este comportamiento se ilustra en la *Figura 3-11*.

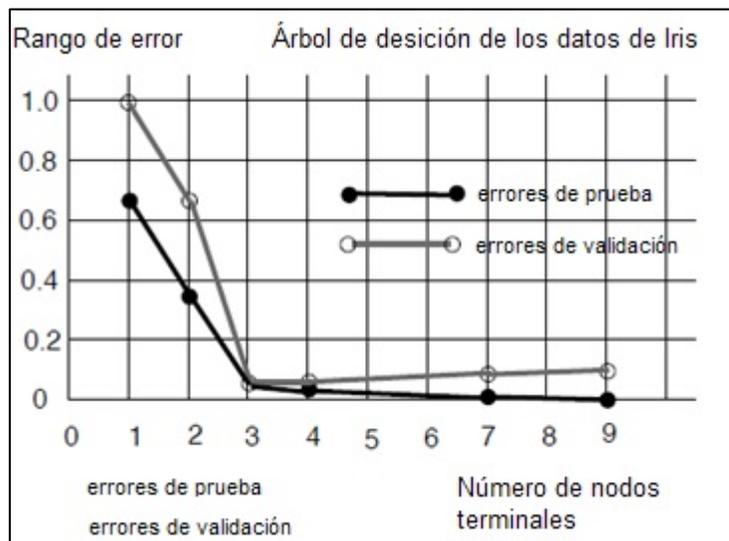


Figura 3-11. Determinando cuando comienza el sobreajuste (Tomado de Weiss y Kulikowski, 1991)

Se puede usar la técnica de validación cruzada para determinar ¿cuándo parar la división de nodos?. Si el error de validación cruzada aumenta, como consecuencia de una división en los nodos, entonces no se debe dividir. Se tiene que tener cuidado sobre cuándo detener la prueba, sin embargo, un ajuste “pobre” o bajo por lo general conduce a más errores en el conjunto de pruebas que tener un sobreajuste.

En lugar de detener el crecimiento de un árbol de decisión también se podría hacer crecer a su tamaño completo y luego “podar” nodos, hojas y ramas hasta que la precisión de la validación cruzada no aumente más. Esta técnica se llama post-poda.

Métodos de longitud de descripción mínima

Una técnica importante de desarrollo de árbol y poda se basa en el principio de la descripción de longitud mínima MDL (*minimal description length*, Rissanen 1978). La idea es que el árbol de decisión más simple que pueda predecir las clases de los patrones de entrenamiento es el mejor árbol.

Considere el problema de transmitir sólo las etiquetas de un conjunto de entrenamiento de patrones, en el supuesto de que el receptor de esta información ya ha ordenado el conjunto de patrones. Si hay m patrones, cada uno marcado por una de las clases R , se podría transmitir una lista de números R con valores m . Suponiendo clases igualmente probables, esta transmisión requeriría bits de $m \cdot \log_2 R$. O, uno podría transmitir un árbol de decisión que etiquete correctamente todos los patrones.

En general, el número de bits (o descripción longitud del mensaje codificado binario) es $t + d$, donde t es la longitud del mensaje requerido para transmitir el árbol, y d es la longitud del mensaje requerido para transmitir las etiquetas de los patrones mal clasificados por el árbol. En un sentido, el árbol asociado con el menor valor de $t + d$ es el mejor o el árbol más económico. El método MDL es una forma de adherirse al principio de la navaja de Occam.

(Quinlan y Rivest, 1989) han propuesto técnicas para la codificación de los árboles de decisión y las listas de etiquetas de excepción y para calcular la longitud de la descripción ($t + d$) de estos árboles y etiquetas. A continuación, utilizan la longitud de la descripción como una medida de la calidad de un árbol de dos maneras:

- a. En el crecimiento de un árbol: utilizan la reducción en la descripción de longitud para seleccionar las pruebas (en lugar de reducción de la incertidumbre).
- b. En la poda de un árbol después de que se haya llegado a error cero: se podan aquellos nodos que logran una disminución de la longitud de la descripción.

Estas técnicas se comparan favorablemente con el método de reducción de la incertidumbre, aunque son bastante sensibles a los esquemas de codificación utilizados

El problema de los sub-árboles

Los árboles de decisión no son el medio más económico de la aplicación de algunas funciones Boleanas. Consideremos, por ejemplo, la función $f = x_1x_2 + x_3x_4$. Un árbol de decisión para esta función se muestra en la *Figura 3-12*. Observe los subárboles mostrados en círculos. La forma de DNF equivalente a la función implementada por este árbol de decisión es $f = x_1x_2 + x_1\bar{x}_2x_3x_4 + \bar{x}_1x_3x_4$. Esta forma DNF no es mínima (en el número de separaciones) y es equivalente a $f = x_1x_2 + x_3x_4$.

La réplica en este árbol significa que se necesita más tiempo para aprender y que los subárboles replicados deben ser aprendidos usando un subconjunto de entrenamiento menor. Este problema a veces se llama problema de la fragmentación.

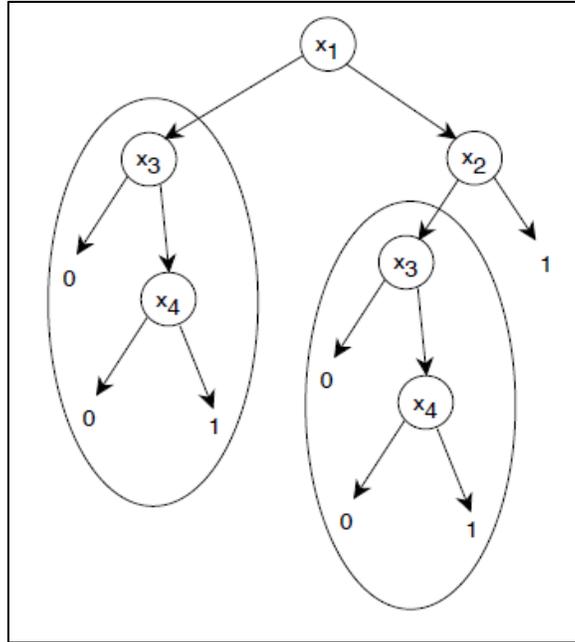


Figura 3-12. Un árbol de decisión con subárboles replicados
(Tomado de Nilsson, 1998)

Varios enfoques pueden ser sugeridos para hacer frente a la fragmentación. Uno es tratar de construir un *gráfico de decisión* en lugar de un árbol (Oliver, Dowe, y Wallace, 1992; Kohavi, 1994). Un gráfico de decisión que implementa las mismas decisiones como la del árbol de decisiones de la Figura 3-12 se muestra en la Figura 3-13.

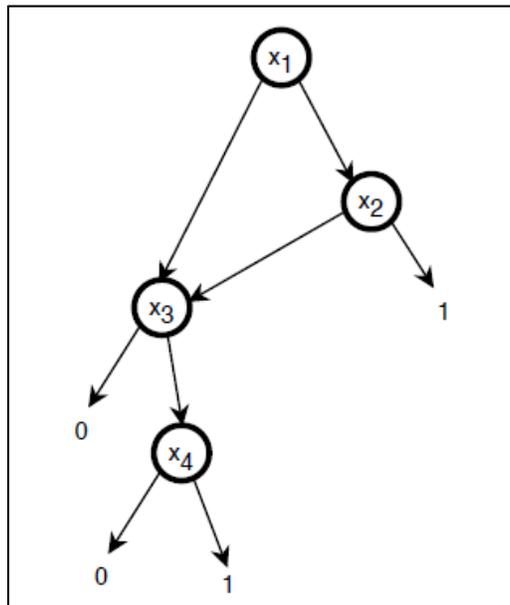
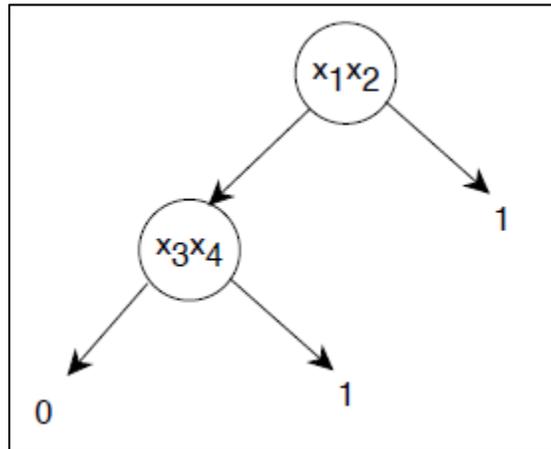


Figura 3-13. Gráfico de decisión
(Tomado de Nilsson, 1998)

Otro enfoque es usar pruebas multivariadas (en lugar de pruebas univariantes en cada nodo). En nuestro ejemplo de aprendizaje $f = x_1x_2 + x_3x_4$, si tuviéramos una prueba para x_1x_2 y una prueba para x_3x_4 , el árbol de decisión podría ser más simplificado, como se muestra en la *Figura 3-14*. Varios investigadores han propuesto técnicas de árboles de decisión de aprendizaje en la que las pruebas en cada nodo son funciones linealmente separables. (John, 1995) da una buena visión general (con varias citas) de aprendizaje de tales árboles discriminantes lineales y presenta un método basado en la “entropía suave.”



*Figura 3-14. Árbol de decisión multivariado.
(Tomado de Nilsson, 1998)*

Un tercer método para tratar con el problema de subárbol replicado implica extraer “reglas” proposicionales del árbol de decisión. Las reglas tienen como antecedentes las conjunciones que llevan hasta los nodos hoja, y como consecuentes el nombre de la clase en el nodo hoja correspondiente a una regla de ejemplo del árbol con el subárbol de repetición de nuestro ejemplo sería: $x_1 \wedge \neg x_2 \wedge x_3 \wedge x_4 \supset 1$. (Quinlan, 1987) discute los métodos para reducir un conjunto de reglas a una simple fijada por: 1) la eliminación del antecedente de cada regla de cualquier conjunto “innecesario”, y luego 2) eliminar reglas innecesarias. Un conjunto de reglas es determinado como innecesario si su eliminación tiene poco efecto sobre la exactitud de la clasificación.

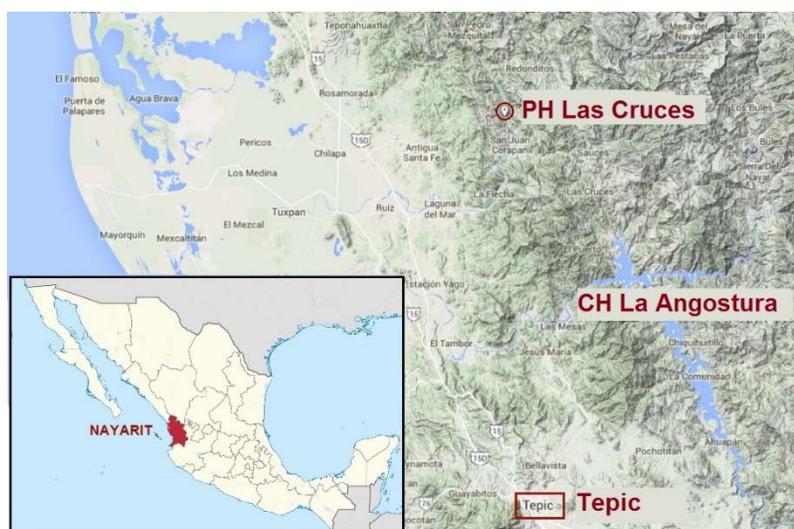
4. CAPÍTULO IV. APLICACIÓN DEL MÉTODO EN UN CASO REAL: SELECCIÓN DE MEZCLAS DE CCR PARA EL P.H. LAS CRUCES

4.1. Antecedentes

Para ejemplificar las ventajas del uso de la minería de datos en el diseño de mezclas de concreto, se presenta el caso del estudio del CCR en el cuerpo de la cortina de un proyecto hidroeléctrico (P.H. Las Cruces). El proceso de selección de mezclas es sin duda una labor en la que las destrezas del diseñador se ponen a prueba ya que debe equilibrar las demandas técnicas de la obra civil (propiedades mecánicas del concreto) con limitantes económicas, sociales e incluso políticas. Se usa este proyecto para describir i) el uso de la metodología, ii) los criterios para simular mezclas (parametrización más adecuada) y iii) la interpretación, desde el punto de vista de un ingeniero, de los resultados obtenidos.

Descripción del sujeto de análisis: Proyecto Hidroeléctrico Las Cruces.

El proyecto sujeto de estudio denominado Proyecto Hidroeléctrico Las Cruces forma parte del Sistema Hidrológico San Pedro-Acajoneta, el cual participa junto con otros 11 proyectos más en este sistema y pertenece al conjunto de los cinco dispuestos sobre el río San Pedro, con un potencial de generación de 1 575 MW. Las Cruces se localiza en la porción central del estado de Nayarit y en la parte baja de la cuenca del río San Pedro, a 65 km y a 30° al noroeste de la ciudad de Tepic (ver *Figura 4-1*).



*Figura 4-1. Localización del proyecto
(Modificado de CFE/GEIC, 2012))*

En el sitio Las Cruces se pretende construir una cortina de 185 m de altura (aunque las dimensiones han sido actualizadas constantemente esta es la que se tomó en los análisis numéricos motivo de esta investigación) la cual permitirá el aprovechamiento hídrico del río San Pedro mediante dos unidades de potencia con una generación media anual de 803 GWh. En el esquema actual, la cortina de 185 m de altura cuenta con una corona en la elevación 245 msnm. La sección de la presa sujeto de estudio se muestra en la *Figura 4-2*.

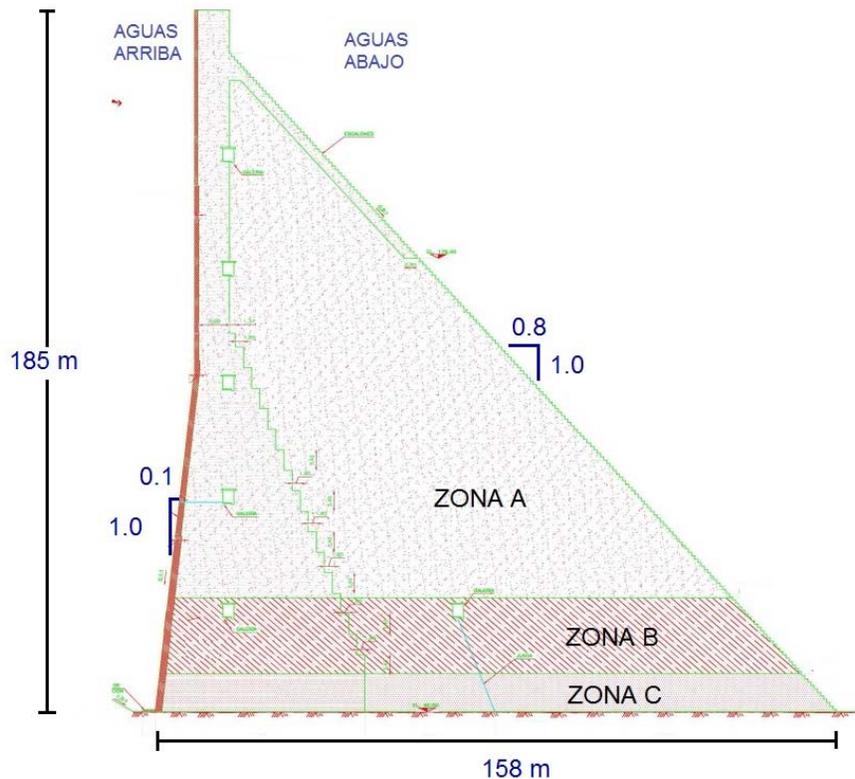


Figura 4-2. Sección y zonificación de la cortina (Modificado de LI, 2015).

En la *Figura 4-2* se muestran las zonas A, B y C que corresponden a mezclas que demandan parámetros mecánicos diferentes, especificados por el diseñador con base en los análisis estático y dinámico de la cortina. El objetivo del diseño de mezclas es encontrar las combinaciones paramétricas que cumplan con los requerimientos específicos, en el caso analizado en este trabajo de investigación, únicamente se trata la demanda de propiedades mecánicas aunque lo aquí expuesto puede ser extendido a la optimización de los procesos, de los costos y de los tiempos.

La minería de datos permite optimizar el proceso en el que se seleccionan los componentes y su grado de intervención en cada mezcla a partir de una restricción de propiedad mecánica, siendo una de sus cualidades más sobresalientes permitir tomar decisiones en escenarios de simulación multidimensional. Este aspecto es particularmente ventajoso en el caso de estudio pues se requiere estudiar las características de cada zona de manera “paralela”. Una herramienta en la que las

variables que definen a cada mezcla (entradas) puedan ser modificadas en rangos no ensayados en laboratorio para verificar el impacto que estos cambios tienen sobre el cumplimiento de alguna restricción (salida) –por ejemplo, un cierto nivel de resistencia – evitar “dar vueltas en círculos” y vuelve eficiente la tarea en el laboratorio y en el gabinete a la vez que permite hacer modificaciones “en línea” en campo cuando existan desviaciones entre lo proyectado y la situación del entorno natural.

4.2. Expresión numérica del problema

La optimización de la selección de mezclas de CCR para presas de gran altura, se utilizaron los fundamentos, técnicas y herramientas de la minería de datos y el aprendizaje máquina. Para ejemplificar el método, se tomó como base el estudio de mezclas de CCR desarrollado para el PH Las Cruces en Nayarit por la Comisión Federal de Electricidad CFE. Los resultados de una acotada batería de pruebas fueron utilizados para construir la base de datos que alimenta al conjunto de entrenamiento y al de validación. Este es el punto de partida en la aplicación de un algoritmo de minería de datos y es la parte más importante en el desarrollo de un modelo correcto: si la expresión numérica del fenómeno por modelar no es la adecuada, el modelo generará números pero éstos representarán situaciones incongruentes.

Teniendo en consideración el riesgo de tomar datos de entrada→salida que no estén bien definidos, se seleccionaron cuidadosamente las parejas que expresan condiciones de mezcla que importan o interesan al proyecto contra las resistencias y módulos medidos en laboratorio que tiene que ver con las restricciones de diseño. El criterio de selección es amplio e intenta almacenar la mayor cantidad de ejemplos de comportamiento de tal forma que se conforme una amplia y útil base de datos pero sólo se usan los patrones (patrón se usa aquí como sinónimo del singular de ejemplo o de prueba) que se califican como los mejores: sin incoherencias, resultados cerrados, condiciones de medición fijas.

Una vez conformado el conjunto de entrenamiento (80% del total de líneas o patrones de la base de datos) y el conjunto de prueba o validación (20% del total de ejemplos o patrones de la base de datos) se verifican las cualidades de los parámetros que se proponen para describir las condiciones de las mezclas, es decir las entradas del modelo. Las salidas en el caso de la optimización de selección de mezclas son, como se ha dicho en secciones anteriores, los parámetros mecánicos requisitos de diseño. A continuación se describen estas variables desde su naturaleza, rango dinámico y aplicabilidad, sin embargo es importante señalar que esto no significa que sean las únicas ni las mejores maneras de asignar parámetros a una mezcla de CCR. Son las condiciones con las que trabajó el personal de CFE y su exploración respondió a limitantes de todo tipo, la mayoría ligadas a las particularidades del propio PH. Se alerta al lector interesado en desarrollar modelos de este tipo para otros proyectos que tome en cuenta este hecho.

4.2.1. Entradas

(Garduño, 2015)

Las entradas son los componentes utilizados en la producción de diferentes mezclas de CCR. Deben seleccionarse, de entre todas las posibles, las entradas que serán incluidas

en el modelo ya que no todos los parámetros que se registran en el laboratorio se juzgan significativos (útiles para expresar dependencia de los parámetros mecánicos de salida) por alguna de las siguientes razones: i) el rango dinámico (variación de su valor) es muy corto (o nulo) de una mezcla de CCR a otra, y ii) la medición fue vaga o no se pudo determinar el grado de control. A pesar de que algunas variables eliminadas tienen importancia teórica, si tienen alguno de los detrimentos anteriores fueron eliminadas del proceso de construcción del modelo. Para algunas entradas, la decisión de incluirse o no en la base de datos final no fue fácil, una mezcla debe responder a varias condiciones mecánicas y se sabe que estas características no dependen de los mismos componentes del concreto. Una de las ventajas de la minería de datos es que las entradas pueden ser intercambiadas, eliminadas o ponderadas en tiempo real conforme la construcción del modelo avanza.

Entonces, las reglas de comportamiento (relaciones paramétricas) corresponden a los modelos específicos, ya que se alimentan exclusivamente de las entradas activas consideradas, y resultan útiles para estudiar a este universo. Sin embargo, esto no significa que las conclusiones generales, los hallazgos físicos, incluso la metodología propuesta pierdan interés para la comunidad científica que se encarga de estudiar a las mezclas de concreto.

Finalmente, las entradas consideradas en este análisis son:

- a) **Tipo de cemento** (expresado en porcentaje, %) Variable de clase que describe el tipo de cemento utilizado en cada mezcla elaborada; en las pruebas desarrolladas no se combinaron cementos para generar una misma mezcla. De acuerdo con esta situación la variable podría incluirse como “clase” en la que puede tomar el nombre de alguno de los tres distintos tipos de cemento referidos por CFE. Sin embargo, esta variable se incluyó “abierta”, es decir, se permitieron combinaciones entre tipos de cemento por lo que la entrada es un vector de tres componentes en el que la primera posición corresponde al CPO, la segunda al CPEG y la tercera al CPP, [CPO,CPEG,CPP]. Por ejemplo, una mezcla a la que sólo se le adicionó cemento tipo CPO tiene como entrada numérica [100,0,0] mientras que una en la que sólo se adicionó CPP su entrada sería [0,0,100]. En el caso de mezclas, si por ejemplo se adicionara Portland ordinario en un cincuenta por ciento, Portland con escoria en un quince por ciento y Portland puzolánico en un treinta y cinco por ciento la entrada se escribiría: [50,15,35].
 - **CPO** - Cemento Portland Ordinario
 - **CPEG** – Cemento Portland con Escoria Granulada de Alto Horno
 - **CPP** – Cemento Portland Puzolánico
- b) **Origen de los agregados** (expresado en porcentaje, %). Se utilizaron dos tipos de agregados de origen diferente, uno es producto de la trituración de rocas sanas y el otro se obtuvo de la recolección del agregado existente en los playones (aluvión). Ambos agregados fueron aprobados para fabricar CCR de acuerdo con su caracterización físico-química.

En este caso sí se probaron mezclas con combinaciones de agregados de distintos orígenes por lo que, al igual que el tipo de cemento, esta es una variable

“abierta” que permite el ingreso en porcentaje de alguno de los orígenes: [CP,SP]. Por ejemplo una mezcla en la que se incluyó la mitad de roca triturada y la mitad de aluvión tiene como vector de entrada [50,50], mientras que una con el ochenta por ciento de aluvión y el veinte por ciento de roca triturada su vector de entrada se expresaría [20,80].

- Cerro Partido (roca triturada), **CP**
 - San Pedro (aluvión), **SP**
- c) **Agua “libre”, w (lts).** es el agua que se agrega a la mezcla con la intención de que reaccione con el cemento, el valor reportado es el que se agregaría por 1 m³ de concreto.
- d) **Contenido de cemento, c (kg).** Es la cantidad de cemento agregado a cada mezcla de concreto, el valor reportado es el que se agregaría por 1 m³ de concreto.
- e) **Contenido de ceniza volante, $f.a.$ (kg).** Es la cantidad de ceniza volante agregada a cada mezcla de concreto, el valor reportado es el que se agregaría por 1 m³ de concreto.
- f) **Relación agua/contenido cementicio, $w/c.c.$** (cemento + ceniza volante). Como principio práctico en la construcción de minería de datos, no se deben agregar parámetros que provengan de la interacción entre otras entradas del mismo modelo, ya que esto podría generar darle más peso o importancia “artificial” a estas variables. Sin embargo para el concreto es sabido que la relación agua/contenido cementicio es un factor muy importante en el desarrollo de las propiedades de una mezcla, con gran significado lingüístico y operacional entre los profesionales involucrados en la construcción con CCR y por esto se decidió sobre su inclusión. En el proceso de construcción y verificación de capacidad de generalización esta variable se vigiló cercanamente para evitar la contaminación de las relaciones paramétricas descubiertas.
- g) **Fracciones de agregado** (expresada en porcentaje, %). El agregado se divide en varias fracciones (recomendadas por los consultores del proyecto), para tener mayor control sobre la granulometría de la mezcla y obtener los beneficios de contar con una buena granulometría. En un principio se pensó en omitir estos valores, pero algunos ejercicios iniciales mostraron el profundo impacto que tienen sobre los valores de salida.

Las fracciones se presentan en vector de cinco componentes en los que la posición ocupada significa un rango en porcentaje y éstos corresponden al tamaño de las partículas de la siguiente manera:

- **Fracción I:** 25-50 mm
- **Fracción II:** 9.5-25 mm
- **Fracción III:** 2.38-9.5 mm

- **Fracción IV:** 0.6-2.38 mm
- **Fracción V:** 0.075-0.6 mm
- **Finos** <0.075 mm

4.2.2. Salidas

(Garduño, 2015)

- a) Resistencia a la compresión, determinada en especímenes sometidos a curado acelerado, **$f_c C/A$ (MPa)**
- b) Resistencia a compresión a los 28 días de edad, determinada en especímenes sometidos a curado estándar, **$f_c 28d$ (MPa)**
- c) Módulo de elasticidad al 40% de la f_c última, determinado en especímenes sometidos a curado acelerado, **$E_{40\%} C/A$ (GPa)**
- d) Módulo de elasticidad al 40% de la f_c última a los 28 días de edad, determinado en especímenes sometidos a curado estándar, **$E_{40\%} 28d$ (GPa)**
- e) Resistencia a tensión indirecta (prueba brasileña), determinada en especímenes sometidos a curado acelerado, **$ft C/A$ (MPa)**

4.3. Rangos dinámicos y distribución de las entradas y salidas

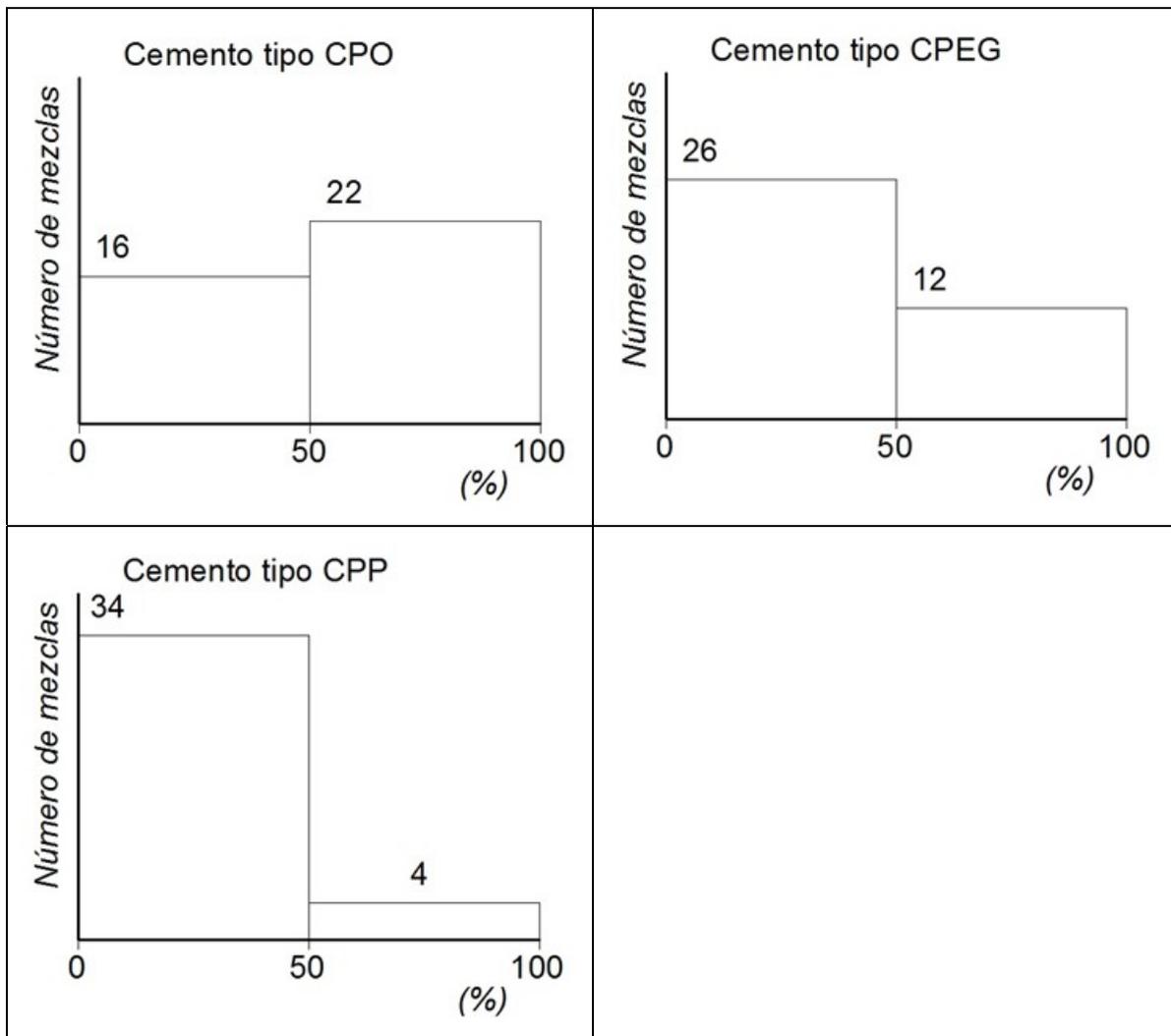
En la elaboración de las recomendaciones de uso del modelo y en el control de las tendencias de las relaciones paramétricas, es necesario conocer, para cada entrada y salida, sus valores mínimos, máximos, la media y su desviación estándar. Con estos valores se establece un puente entre el modelo matemático (minería) y el fenómeno real que está siendo explorado en el laboratorio. La abstracción simplificada de la *realidad* genera obligadamente una cierta discrepancia entre lo que se observa y lo previsto por el modelo pero a través del conocimiento de los límites (mín-máx) o fronteras de aplicación y la representatividad de los comportamientos se pueden evaluar y juzgar estas diferencias entre la realidad y la teoría.

Entradas

- a) Tipo de cemento

Al tratarse de una variable de clase, en la que en porcentaje se indica la cantidad del tipo de cemento adicionado a cada mezcla ensayada, el análisis de máximos y mínimos resulta un tanto árido pues el rango lógico es de 0 a 100. Lo que resulta interesante es observar la tendencia de los patrones, mayor número de mezclas con CPO y CPEG fueron ensayadas (en equilibrado porcentaje) mientras que CPP fue la condición que menos fue experimentada.

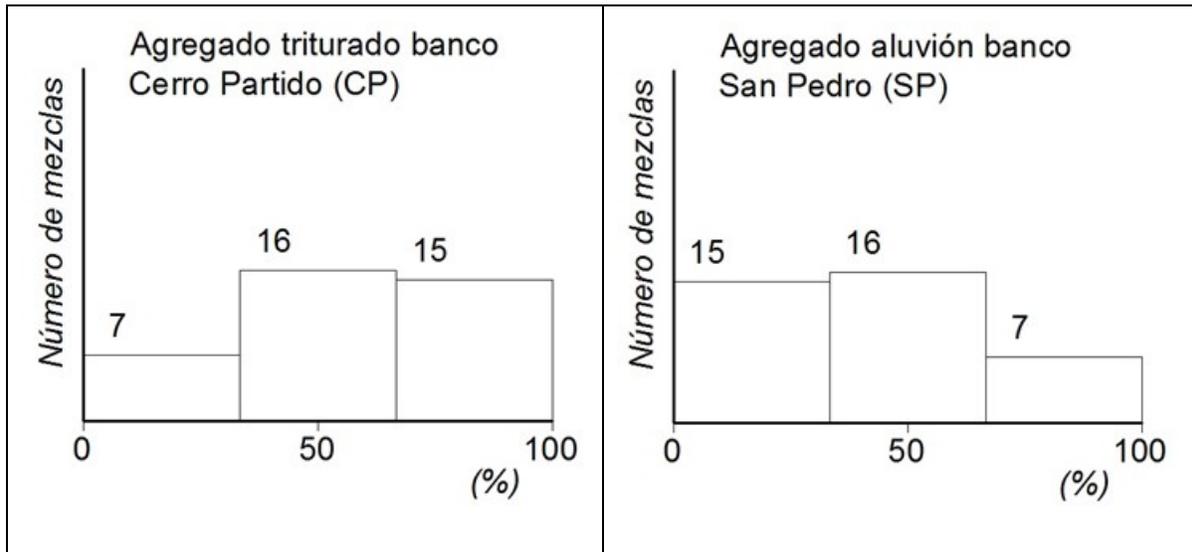
	CPO	CPEG	CPP
Mínimo, %	0	0	0
Máximo, %	100	100	100
Media, %	57.895	31.579	10.526
Desv. Std.	50.036	47.107	31.101



b) Origen de los agregados

Esta es una variable de clase en la que el porcentaje indica la cantidad de material de distinto origen que fue agregado a la mezcla. Como en el caso del tipo de cemento adicionado a cada mezcla ensayada, el análisis de máximos y mínimos genera sólo análisis entre 0 y 100. En el análisis de disposición de los patrones se observa que se ensayaron un mayor número de mezclas con materiales de distinto origen y en menor número mezclas de CCR en las que sólo se usara un tipo de agregado.

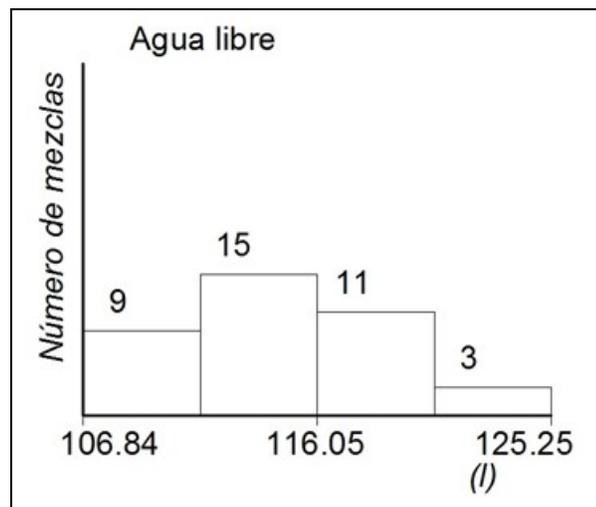
	CP	SP
Mínimo, %	0	0
Máximo, %	100	100
Media, %	60.526	39.474
Desv. Std.	37.05	37.05



c) Agua libre (lts.)

El contenido de agua libre varía en ~20 lts. Más de la mitad de las pruebas se realizaron alrededor del contenido medio y sólo tres fueron probadas adicionando el máximo valor (125 lts.) mientras que en nueve del total de las treinta y ocho se usaron los límites mínimos.

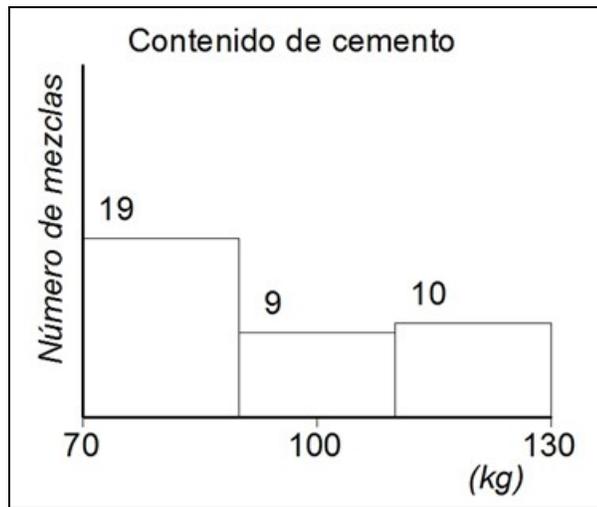
Mínimo, l	106.84
Máximo, l	125.25
Media, l	115.059
Desv. Std.	4.259



d) Contenido de cemento (kg)

Debido al enorme impacto económico que tiene el contenido de cemento sobre la obra civil, la CFE ensayó en un mayor rango de variación los contenidos de las mezclas. Partiendo de 70 kg y hasta 130 kg fueron ensayadas muestras en proporciones convenientes para los fines del proyecto. Observe que se prueban más probetas cuyo contenido ronda los valores mínimos, lo cual es lógico y sigue la restricción económica.

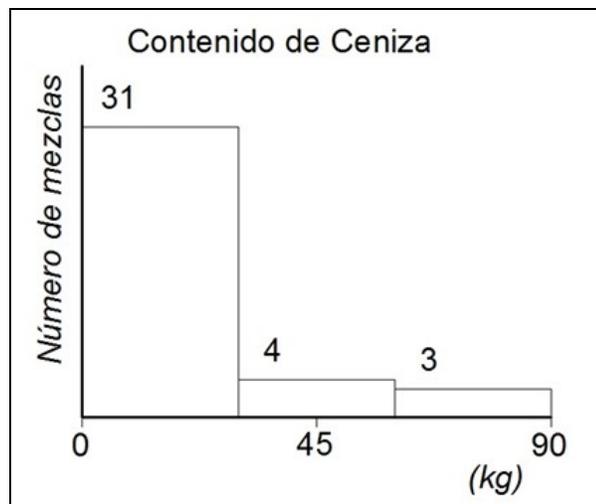
Mínimo, kg	70
Máximo, kg	130
Media, kg	102.632
Desv. Std.	20.226



e) Contenido de ceniza volante (kg)

El número de kilogramos que fueron adicionados a las mezclas arranca en no inclusión y llega hasta los 90 kilos. Esta variable es un ejemplo claro de los casos de parámetros que a pesar de haber sido probadas en un rango de variación importante, su repercusión en el modelo está sesgada debido al número de pruebas que fueron ejecutadas entre el valor mínimo y la media y ésta y el valor máximo. Como se observa en la gráfica de barras, la mayoría de las pruebas se realizaron con contenidos que están por debajo de la media en proporciones menores.

Mínimo, kg	0
Máximo, kg	90
Media, kg	12.895
Desv. Std.	25.828



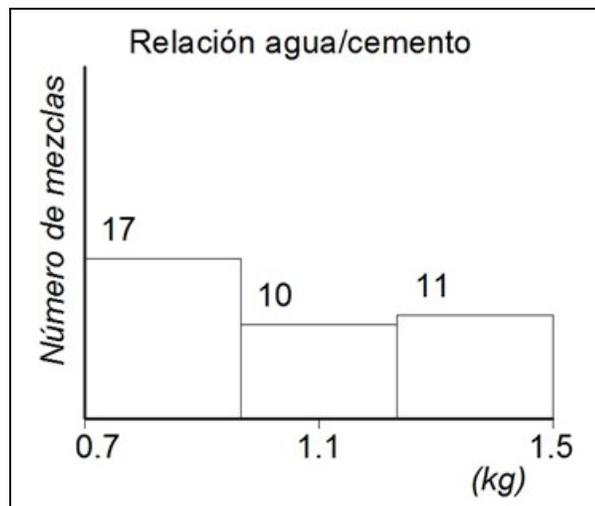
f) Relación agua/cemento w/c

La relación agua/cemento es uno de los parámetros más importantes de la composición del concreto. Tiene influencia sobre la resistencia, la durabilidad, la fluencia y la retracción. La relación w/c es un valor característico ampliamente usado, entendido y relacionado en las tecnologías del concreto. Ya que determina la estructura interna de la pasta de cemento endurecida es de fundamental importancia y se espera que cualquier modelo de análisis la pondere sobre otros factores menores.

La relación w/c crece cuando aumenta la cantidad de agua y decrece cuando aumenta el contenido de cemento y en todos los casos cuanto más baja es la relación más favorables son las propiedades de la pasta de cemento endurecida.

La equilibrada distribución de mezclas ensayadas con valores diferentes de w/c en casi el mismo número de probetas es calificada positivamente.

Mínimo	0.7
Máximo	1.5
Media	1.034
Desv. Std.	0.245

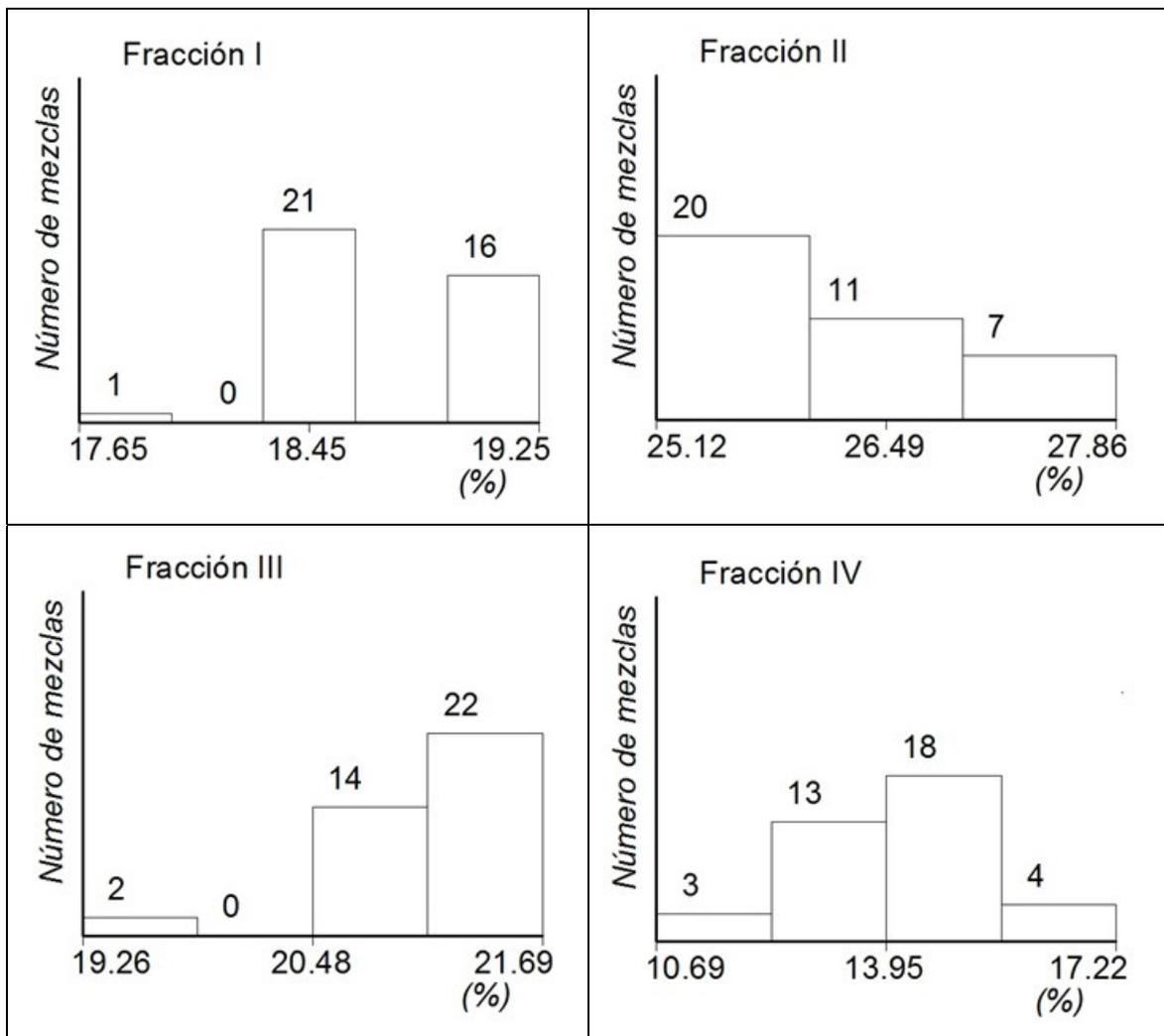


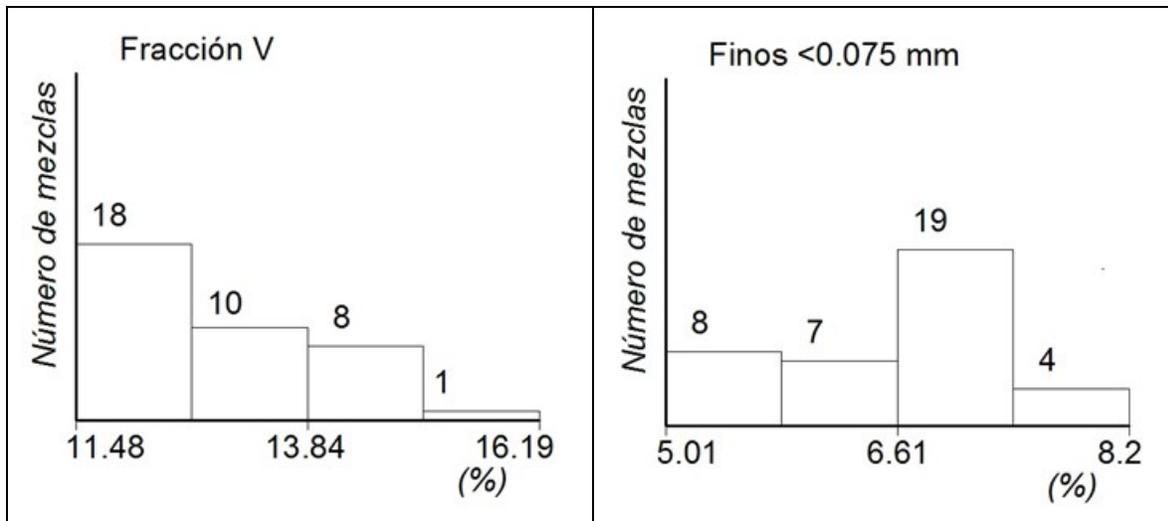
g) Fracciones de agregados (expresados en porcentaje, %)

La distribución de los tamaños de las partículas o granulometría de un agregado es una característica importante debido a que determina los requerimientos de la pasta para lograr un concreto trabajable. Como ya se mencionó, ya que el cemento es el componente más costoso del concreto, es deseable, minimizar el costo del concreto utilizando la menor cantidad de pasta de tal forma que se logre un concreto que pueda ser manejado, compactado, acabado y proporcionar la resistencia y durabilidad necesarias. La distribución de tamaño de partículas se refiere al ensamble de agregados ligeramente compactados que se mantendrán unidos a través de la pasta de cemento por lo que la cantidad de pasta depende de la cantidad de espacios vacíos que deben ser llenados y de la cantidad total de superficie de los agregados que debe ser recubierta por la pasta.

El personal de la CFE definió muy oportunamente la batería de pruebas tomando en cuenta lo anterior y probó combinaciones que intentaron disminuir el volumen de los espacios entre las partículas. Las mezclas, de acuerdo con lo que se observa en la gráfica de barras, contienen rango de tamaños en los que las partículas más pequeñas se adicionan para ocupar los espacios vacíos dejados por las partículas más grandes. Esto disminuye los espacios vacíos y se consiguen los requerimientos de pasta por un lado y el aumento de resistencia por otro, o por lo menos así se espera.

	I	II	III	IV	V	<0.075mm
Mínimo, %	17.65	25.12	19.26	10.69	11.48	5.01
Máximo, %	19.25	27.86	21.69	17.22	16.19	8.2
Media, %	18.803	26.161	21.018	14.505	13.014	6.592
Desv. Std.	0.348	0.812	0.57	1.487	1.099	0.739





Salidas

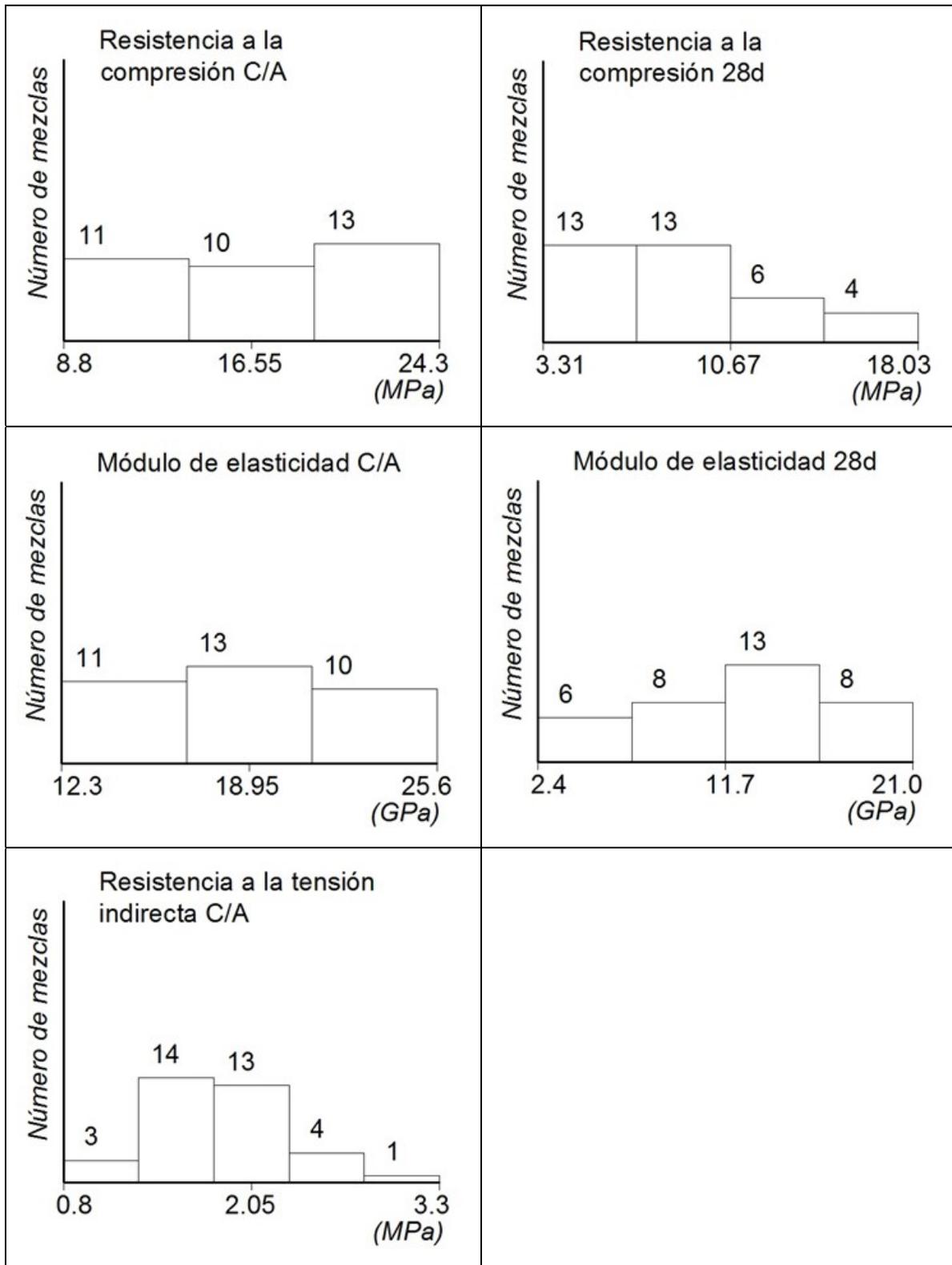
En el diseño del PH se establecieron los parámetros mecánicos que eran necesarios en la(s) mezcla(s) que constituirán la cortina. Con base en una profunda revisión en la literatura y el levantamiento detallado de las experiencias del personal de la CFE se programaron las baterías de pruebas de tal forma que el universo de búsqueda (salida) correspondiera a las exigencias de diseño.

Los parámetros mecánicos objetivo, como un todo, dibujan la principal característica estructural del concreto que es la de resistir muy bien los esfuerzos de compresión, y en menor nivel los de tensión, flexión, y cortante. La mezcla de CCR que mejor se ajuste a los requerimientos de diseño mecánicos estará situada en el rango dinámico que se expresa en las siguientes gráficas y tablas. Las mezclas ensayadas están equilibradamente representadas en todo el rango de resistencia a la compresión y módulo de elasticidad, sin embargo los resultados sobre la resistencia a la tensión indirecta, en casi todas las pruebas, están alrededor de la media.

a) Mínimos, máximos, media y desviación estándar.

	<i>f_c C/A</i> MPa	<i>f_c 28d</i> MPa	<i>E_{40%} C/A</i> GPa	<i>E_{40%} 28d</i> GPa	<i>f_t C/A</i> MPa
Mínimo	8.8	3.31	12.3	2.4	0.8
Máximo	24.3	18.03	25.6	21	3.3
Media	16.715	8.864	18.727	12.669	1.865
Desv. Std.	4.648	3.641	4.004	4.631	0.494

b) Distribución de cada salida.



Los mejores datos para construir un modelo de predicción de propiedades son aquellos que están bien distribuidos en las entradas (mismo número de patrones en los deltas convenientes del rango utilizado) cubriendo de manera global el movimiento de las salidas (objetivo). En este trabajo la CFE enfrentó inconvenientes para seguir ensayando situaciones y experimentando en áreas de la zona de búsqueda que fueran de interés práctico así, a pesar de no contar con detalles en la base de datos, se considera suficiente para construir un modelo útil para los requerimientos de diseño.

4.4. Gráficas globales

Una herramienta que auxilia en la comprensión de las relaciones entradas→salidas dentro de la base de datos, y por lo tanto en las predicciones del modelo, son las gráficas globales o multiparamétricas. A continuación se presentan algunos ejemplos de las construidas en este trabajo de investigación.

Los aspectos clave en la lectura de estas gráficas son

- Determine el parámetro al que se ha asignado la variación color
- Seleccione el parámetro que pertenece al eje horizontal
- Seleccione el parámetro que pertenece al eje vertical
- Concluya sobre tendencias entre los ejes, entre el color y uno de los ejes o entre el color y ambos ejes

Lea el siguiente ejemplo. La barra de color corresponde al valor de la relación agua/cemento, los colores fríos (azules) representan los valores más bajos de la relación mientras que los cálidos representan los más altos.

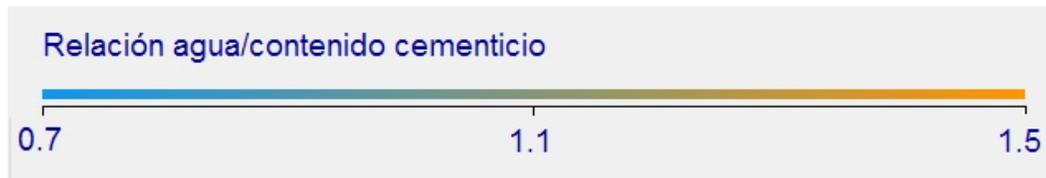
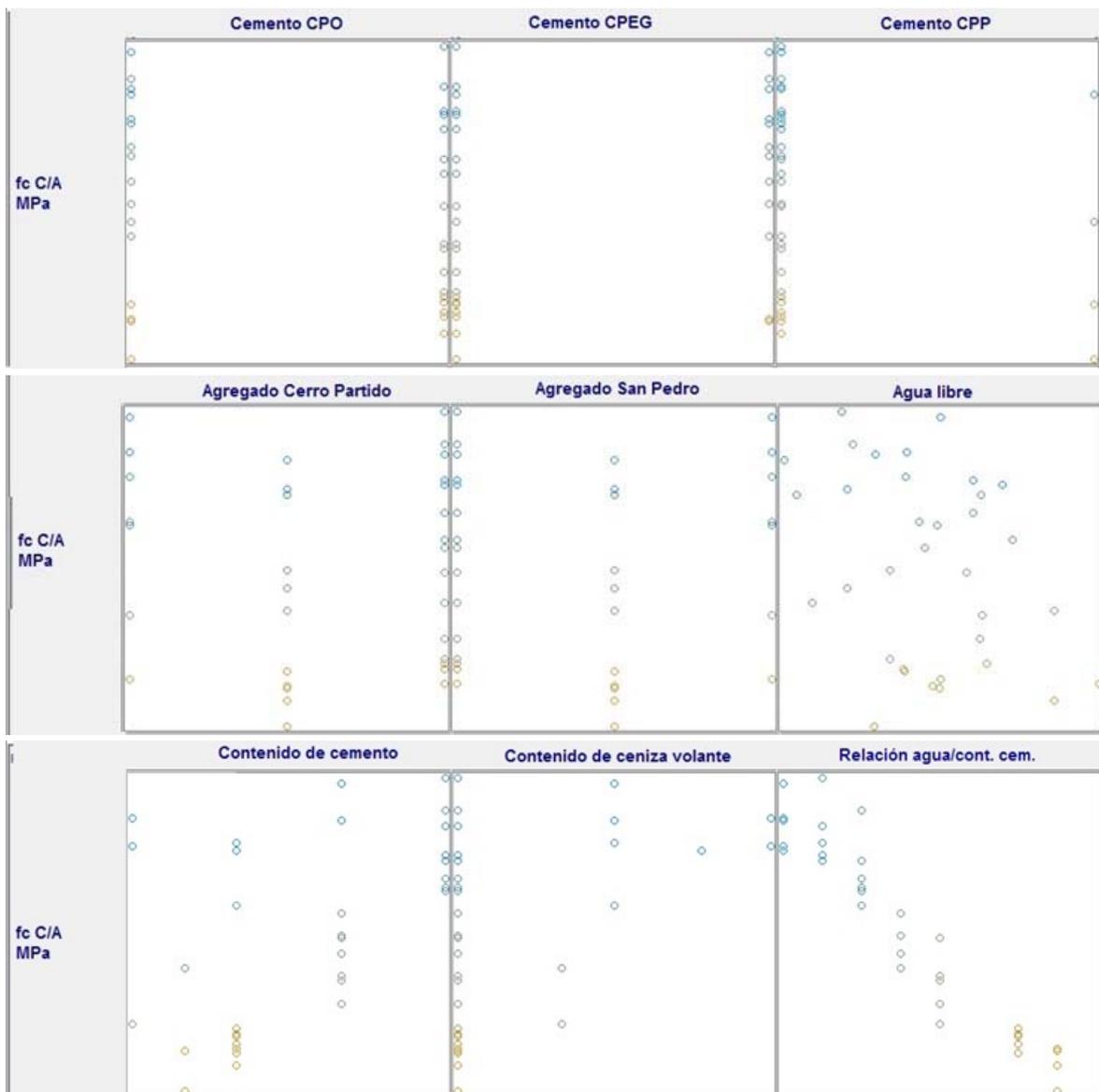


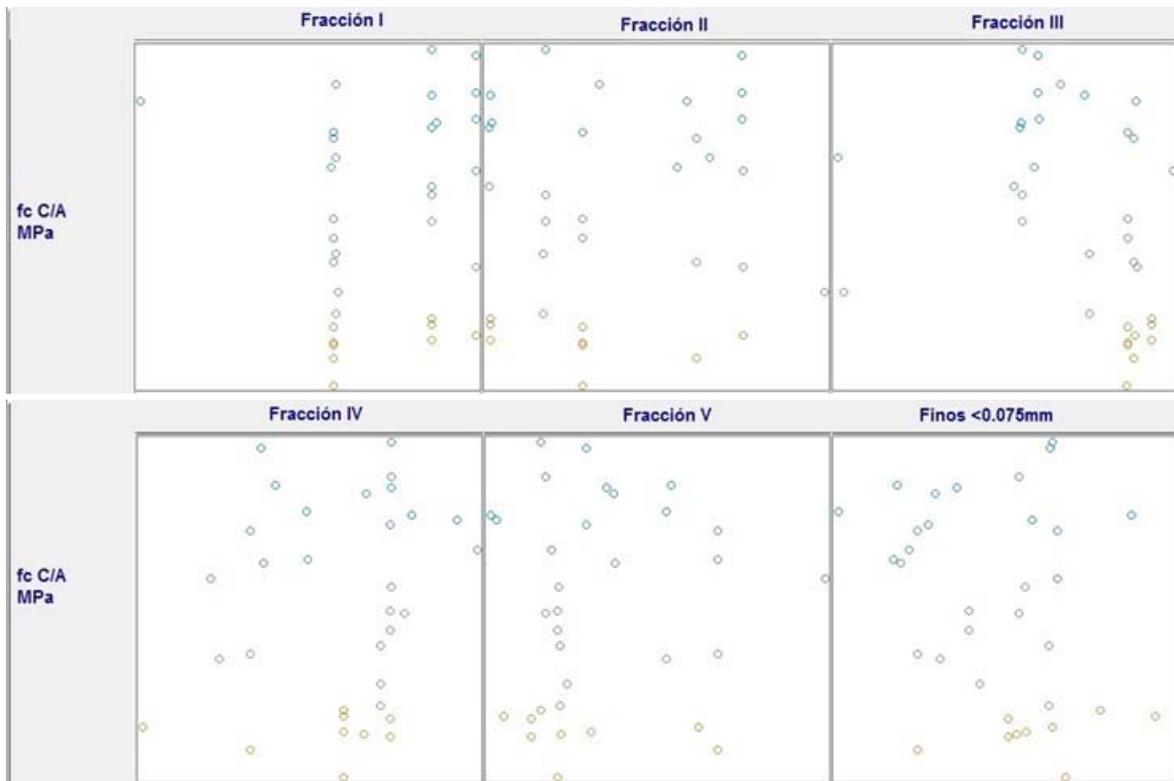
Figura 4-3 Escala de color en función de la relación agua/contenido cementicio

En el eje horizontal se han colocado los valores de las entradas del modelo mientras que en el eje vertical se grafican los datos de resistencia a la compresión obtenidos de las pruebas. Observe los primeros tres recuadros, al tratarse de variables de clase y no haber sido probadas mezclas en combinación de cementos, las posiciones son únicamente extremos: cero o cien. Como es de esperarse, conforme crece la relación w/c el valor de f_c decrece. Sin embargo, si se analiza el gráfico que involucra al agua libre y su relación con w/c y f_c C/A ya no es tan simple ni directo determinar la dependencia ni en qué grado sucede. La primera de las relaciones agua libre *versus* f_c C/A es altamente no lineal, mientras que esto sucede se repite el patrón de encontrar los fríos de w/c en la parte alta del gráfico y los cálidos en la baja.

En cuanto a los agregados obviamente las mezclas 50-50 generan los mismos comportamientos pero los ensayos con material exclusivo de Cerro Partido son los que desarrollaron f_c C/A más interesantes para los fines del proyecto. El análisis de las fracciones es muy interesante, estos sencillos gráficos indican que esta es una variable que puede ser usada como factor que detone resistencias más altas cuando los demás parámetros no deban o puedan ser modificados. Observe cómo algunas fracciones en todo el rango probado generaron resistencias en los límites más altos, las combinaciones entre éstas podrían ser significativas cuando se necesiten porcentajes más altos de f_c C/A.

Resistencia a la compresión en curado acelerado, f_c C/A



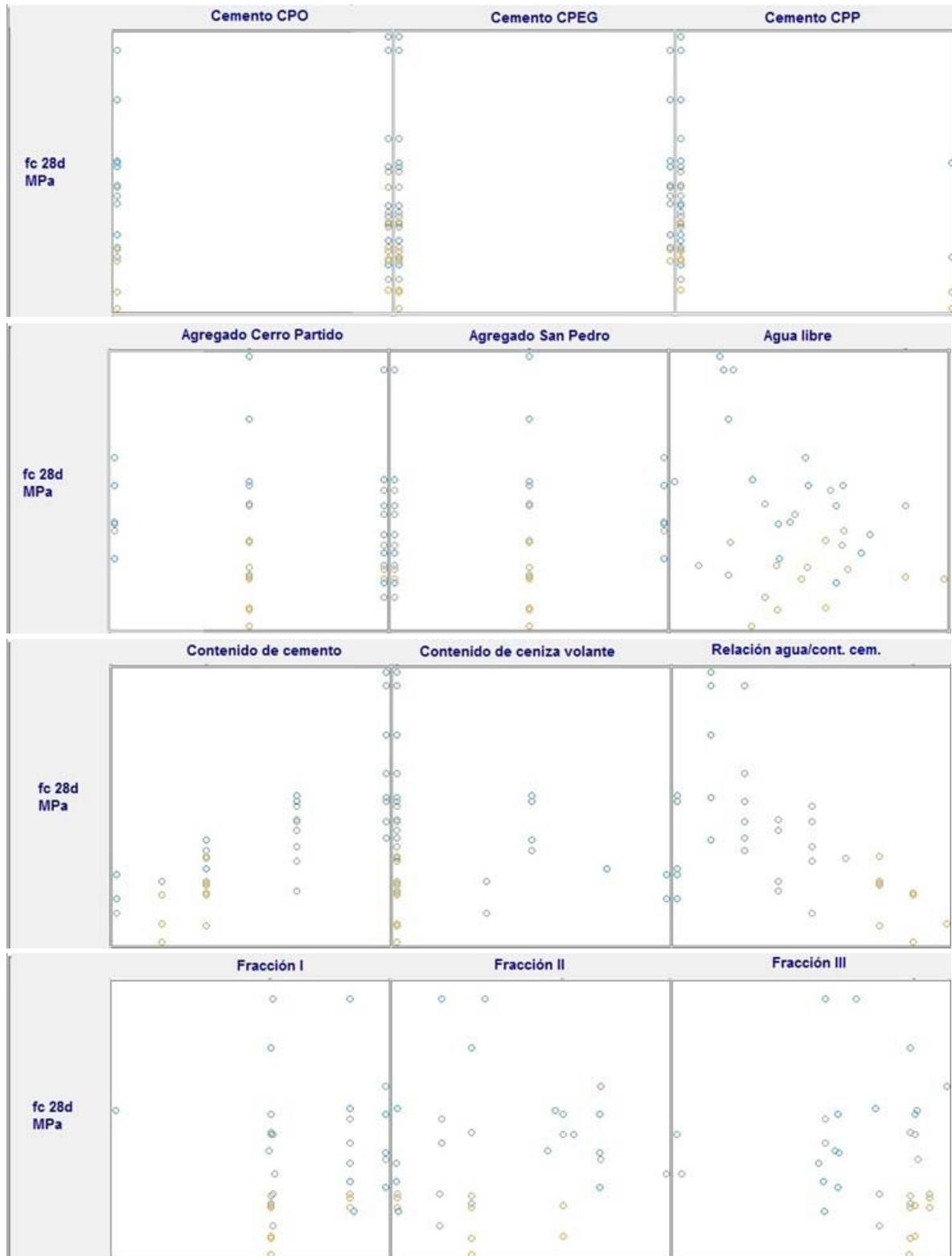


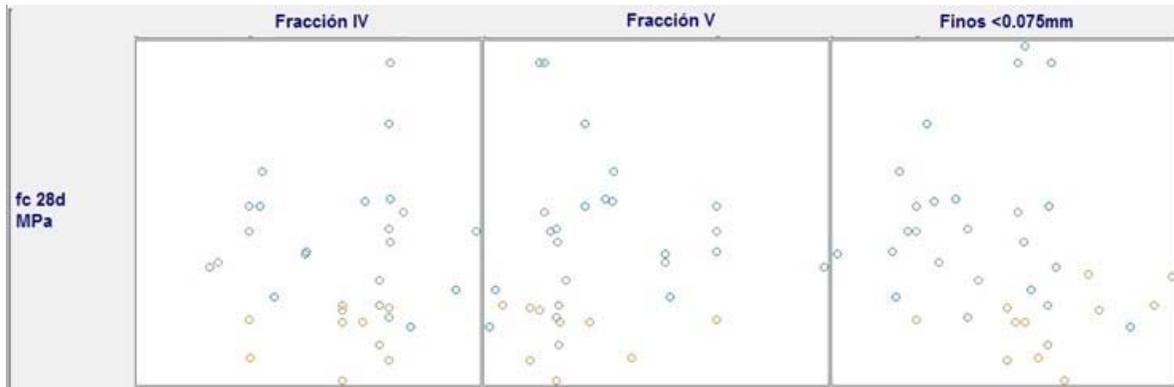
Resistencia a la compresión a 28 días en curado estándar, f_c 28d

Al igual que en el caso de f_c C/A, en el eje horizontal se muestran los valores de las entradas del modelo mientras que en el eje vertical se presentan los datos de resistencia a la compresión, en curado estándar, obtenidos de las pruebas. En los tres primeros recuadros se pueden observar los comportamientos de la variable de clase *tipo de cemento*; al igual que en el ejercicio anterior las posiciones son únicamente extremos (0 o 100) y las conclusiones sobre el impacto en f_c resultan un tanto áridas si no se combinan con otros parámetros. A diferencia del marcado comportamiento lineal del crecimiento de la relación w/c con el decremento del valor de f_c en este caso la tendencia es distinta con un rango de resistencia constante a pesar del ligero crecimiento de la relación w/c . Las relaciones entre agua libre, w/c y f_c no siguen patrones identificables en este tipo de gráficos lo único reconocible es que los fríos de w/c están en la parte alta del gráfico y los cálidos en la baja aunque debe destacarse que no siguen los patrones que marca el f_c C/A, este aspecto puede tener una gran repercusión cuando se eligen los tiempos y los aditivos con los que se intenta cumplir los objetivos de diseño. Existen contenidos de cemento y de ceniza volante óptimos y a pesar de que en los gráficos no se pueden identificar los valores absolutos en los que se consiguen los más altos f_c , sí resulta muy claro que no son los valores más bajos ni los más altos en los contenidos los que dirigen la respuesta en resistencia. Los agregados mantienen los mismos comportamientos y las tendencias con material exclusivo de Cerro Partido. El análisis de las fracciones indica que este vector tiene un profundo impacto en el valor de la resistencia y que es así, como vector, que debe ser estudiado. Esto resulta muy interesante pues ratifica la postura de

Capítulo IV. Aplicación del método en un caso real: Selección de mezclas para el P.H. Las Cruces, Nay.

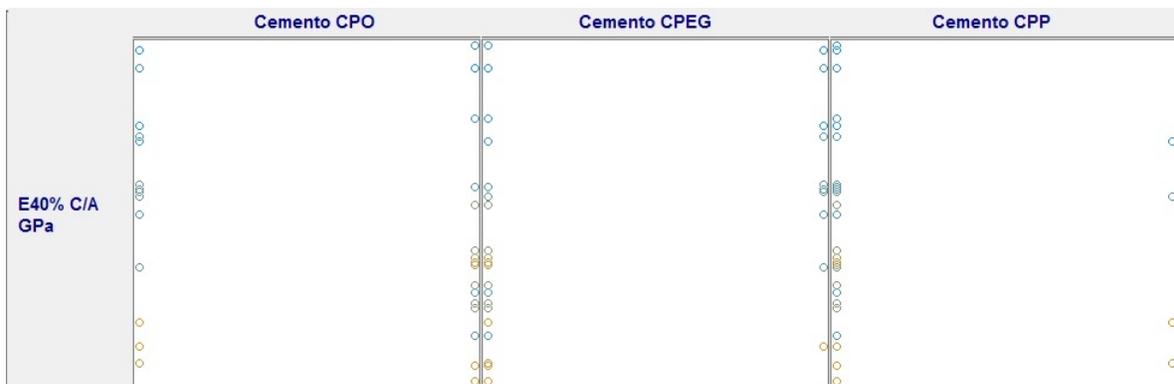
este trabajo de investigación sobre la necesidad de estudiar estos materiales con herramientas altamente dimensionales y no lineales.



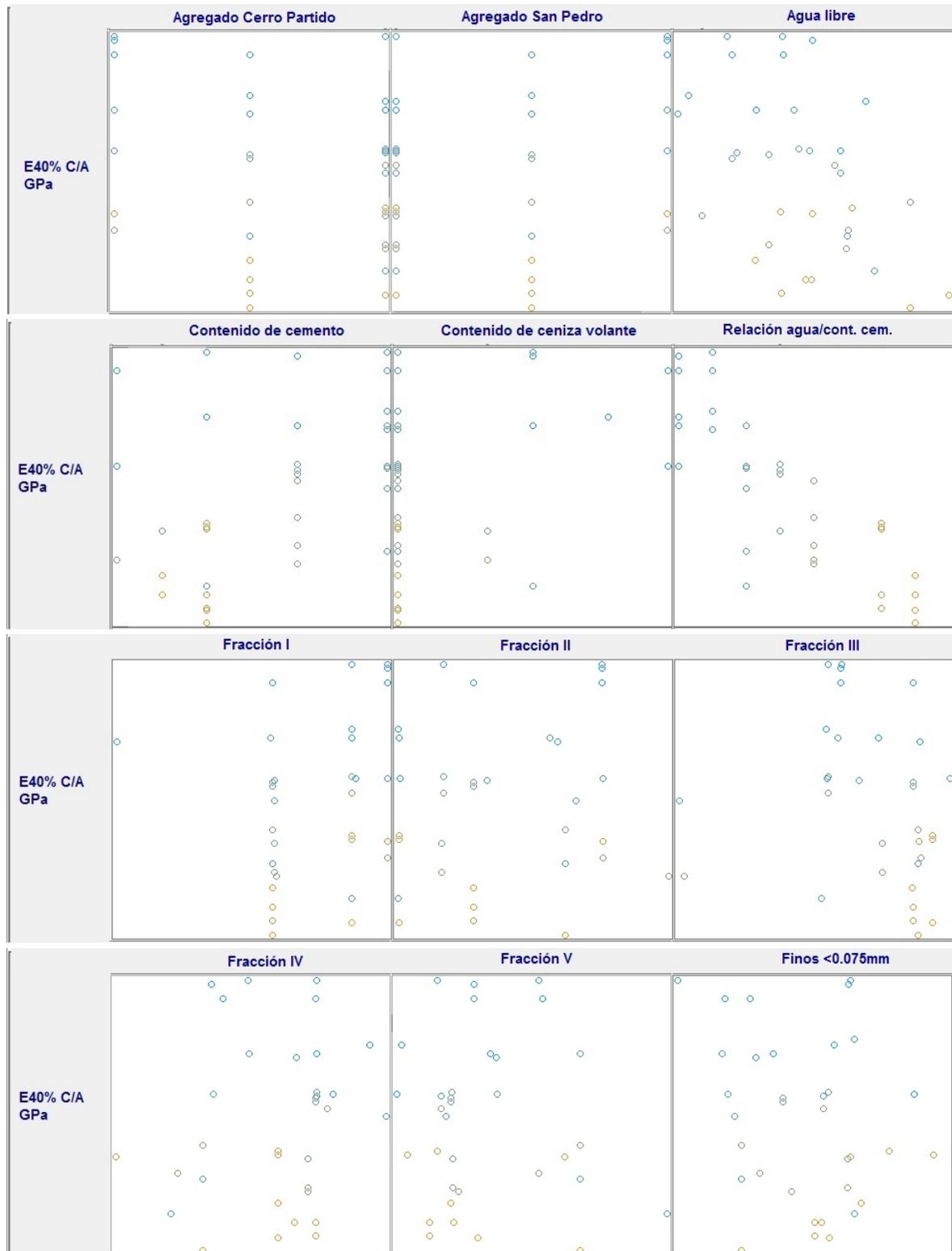


Módulo de elasticidad en curado acelerado, E_{40%} C/A

En cuanto al módulo de elasticidad en curado acelerado, también se muestra en el eje horizontal los valores de las entradas pero ahora en el vertical se grafican los datos de $E_{40\%}$ C/A obtenidos de las pruebas. Mientras que los primeros tres recuadros, al tratarse de variables de clase y no haber sido probadas mezclas en combinación de cementos, son similares los comportamientos observados para f_c , los patrones en los agregados son distintos, ubicando colores cálidos de w/c a media gráfica (valores medios de módulo) lo que significa mucho en cuanto a proporciones ideales para las necesidades de diseño. El *agua libre* y su relación con w/c y $E_{40\%}$ C/A ya no es simple ni directo de determinar, la dependencia ni en qué grado sucede son expuestas en este gráfico 3D, sin embargo se puede señalar que a diferencia de los resultados para f_c , en este caso los valores cálidos de w/c se encuentran en todo el rango de *agua libre* y significan valores altos de módulo, entonces es necesario estudiar en lo multiparamétrico estas mezclas si se quiere llegar a conclusiones correctas y a toma de decisiones eficientes. Para los agregados obviamente las mezclas 50-50 generan los mismos comportamientos y en este caso no se observan a simple vista resultados distintos para los dos tipos ensayados. El análisis de las fracciones es interesante y muestra una sustancial no linealidad con respecto a los patrones mostrados para f_c . De la misma manera que en los ejercicios con las salidas anteriores, estos sencillos gráficos indican que esta es una variable que puede ser usada como factor clave para modificar los valores del módulo en las mezclas elegidas.

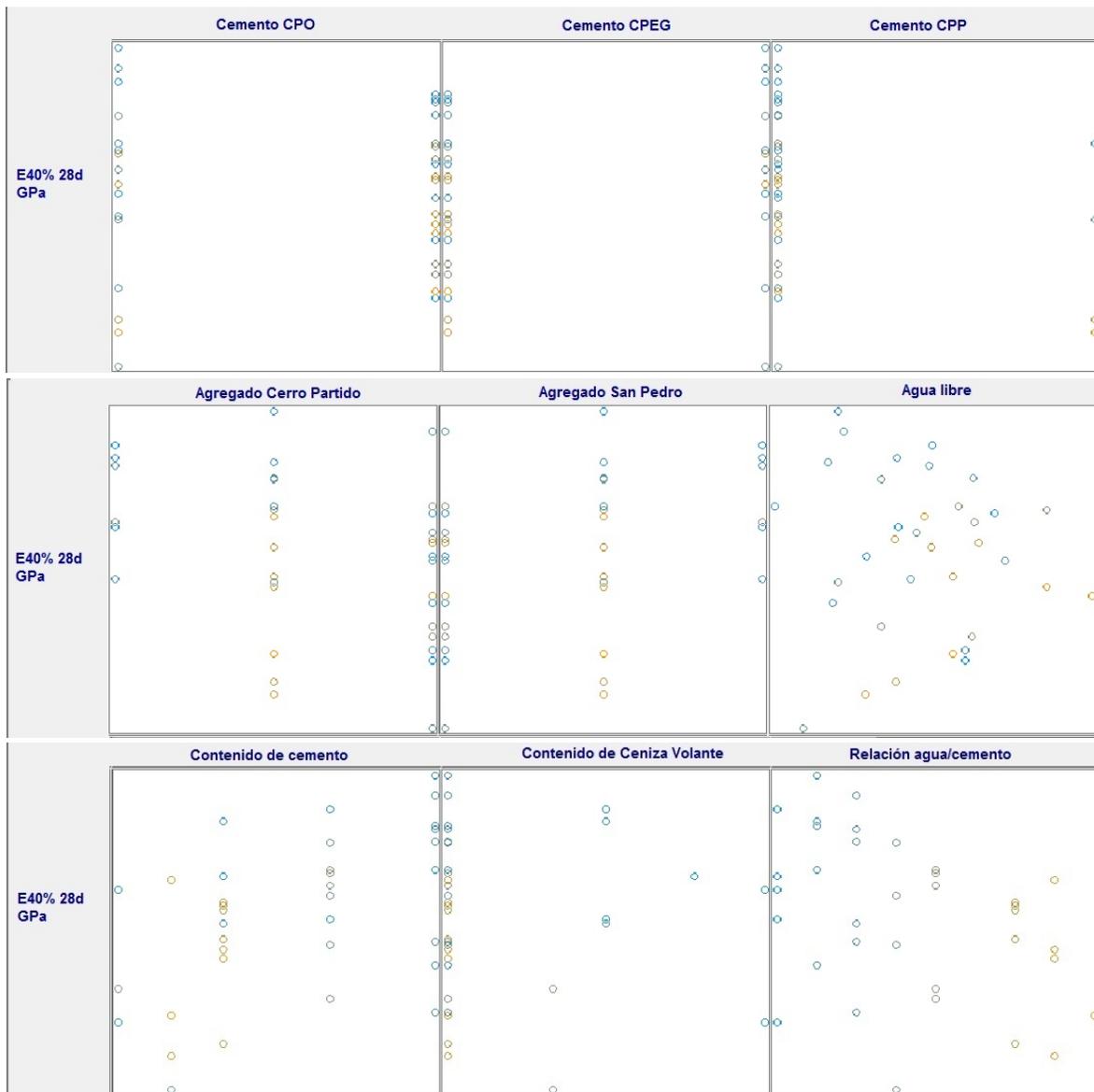


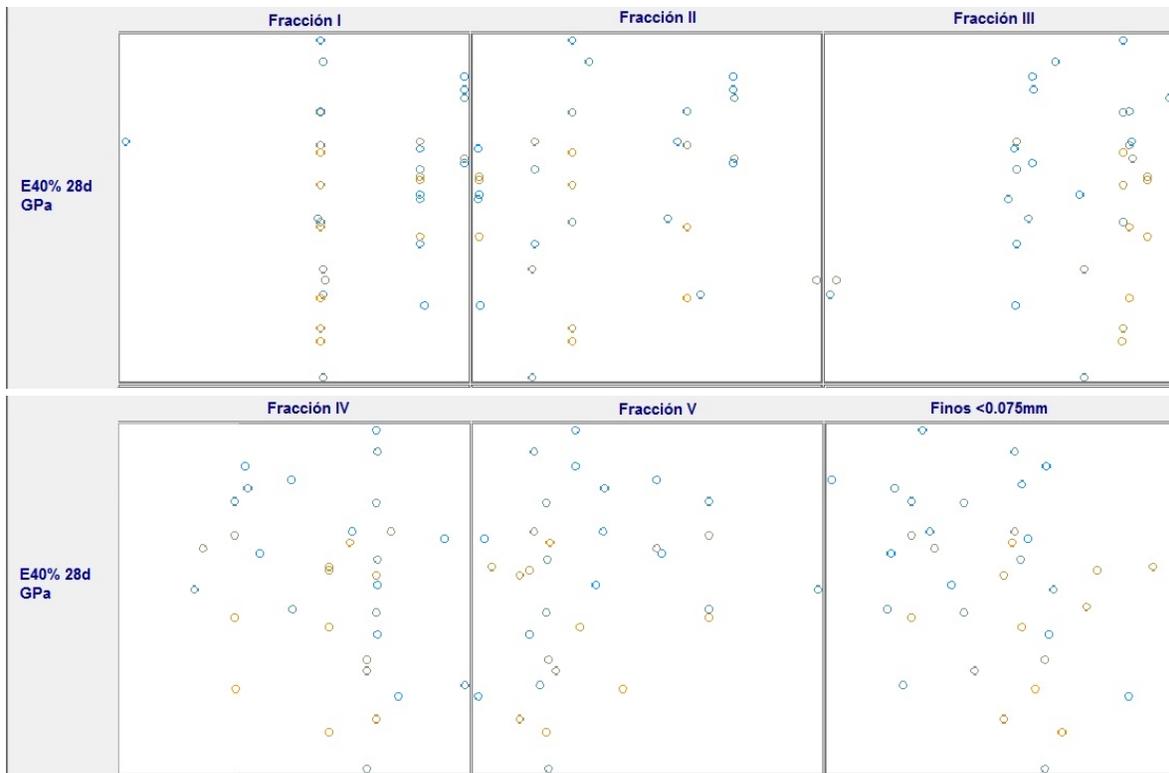
Capítulo IV. Aplicación del método en un caso real: Selección de mezclas para el P.H. Las Cruces, Nay.



Módulo de elasticidad a 28 días en curado estándar, E40% 28d

El módulo de elasticidad en curado estándar se comporta paraméricamente de forma muy similar al de C/A. El agua libre y su relación con w/c muestran tendencias constantes: valores cálidos de w/c se encuentran en todo el rango de agua libre con valores altos de módulo y la alta dimensión es latente si se observan la aparente distribución errática de los puntos muestreados. Los agregados, como totales y como mezclas generan los mismos comportamientos en el caso del módulo C/A y las fracciones mantienen la no linealidad. Este parámetro corrobora la necesidad de tomar muy en cuenta la contribución como vector de las fracciones agregados.

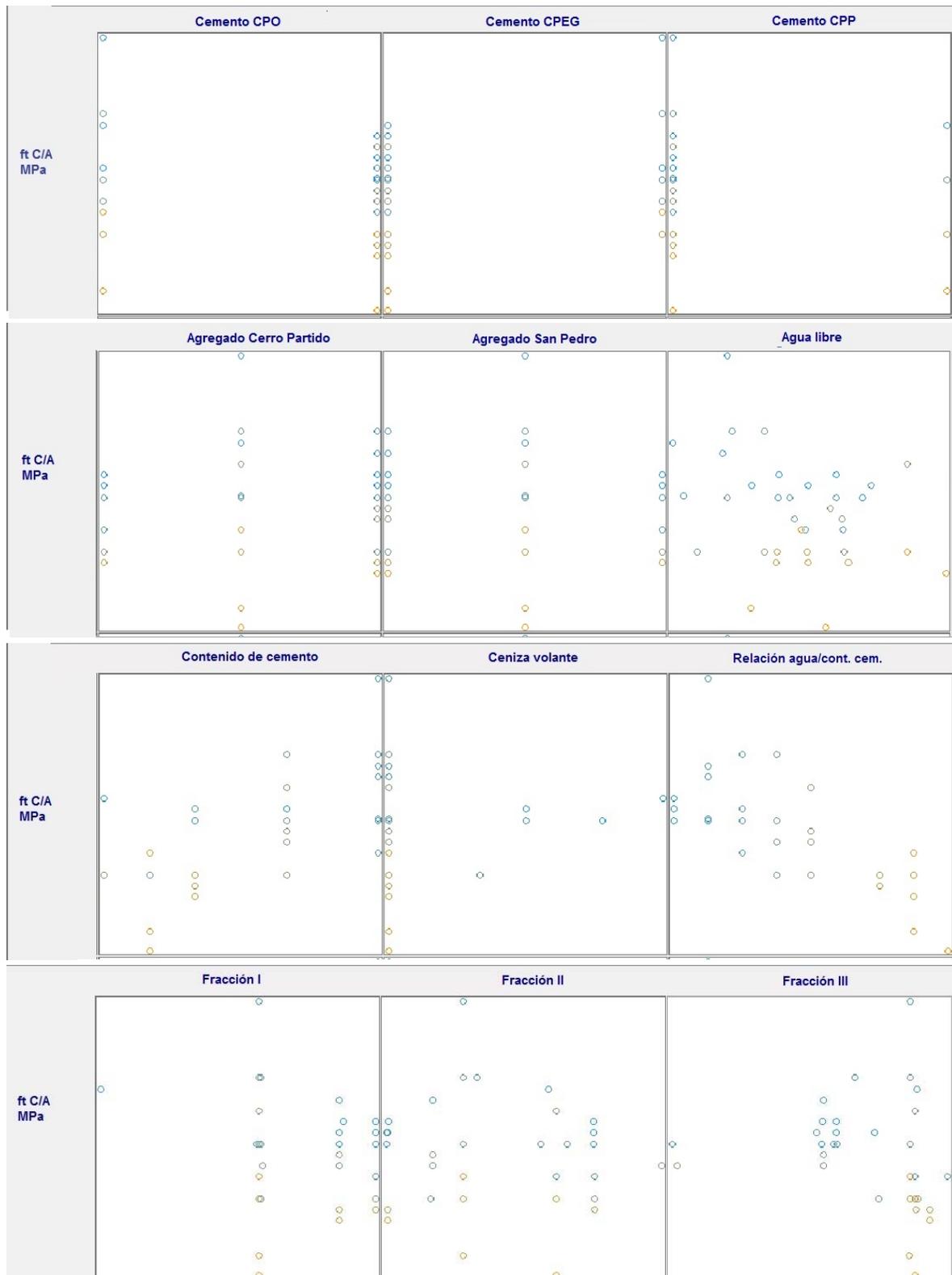


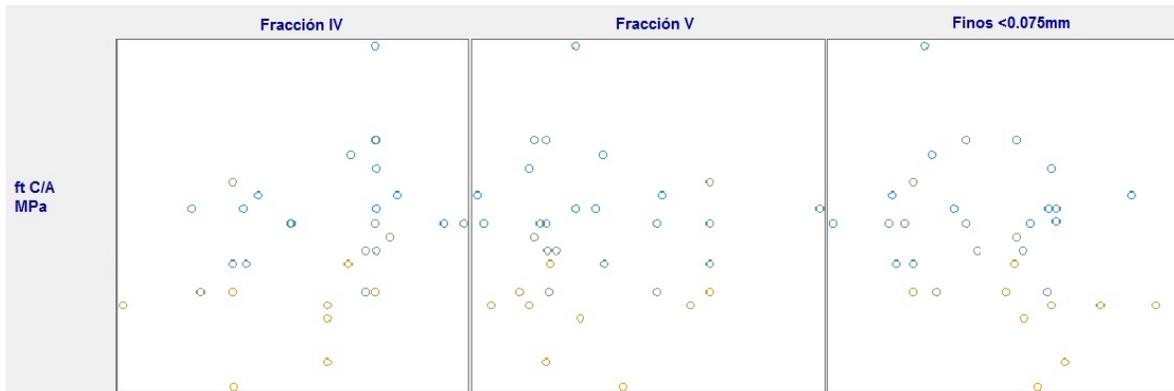


Resistencia a la tensión indirecta en curado acelerado, f_t C/A

Este importante parámetro para calificar a la mezcla se estudió de la misma manera que los módulos y los f_c , así que en el eje horizontal se han colocado los valores de las entradas del modelo mientras que en el eje vertical se grafican los datos de f_t C/A. A diferencia de los patrones anteriores, en el caso de la resistencia al tensión se observa un ligero determinismo en todas las entradas estudiadas. Los rangos de los fríos y los cálidos se ubican conforme a lo que se espera teóricamente. Importante es que en este caso los rangos de f_t se definen claramente en bandas de w/c . No se observan comportamientos lineales claros pero como en el caso del agua libre, los valores más altos de f_t y w/c se localizan en áreas muy definidas del rango en el que se probó esta variable. En cuanto a los agregados las mezclas 50-50 y los ensayos con 100 generan comportamientos coincidentes (altos valores de f_t y w/c) en todo el rango probado lo que hace muy difícil concluir al respecto de su impacto en el valor de la f_t sin tomar otras variables como puntos de referencia. El análisis de las fracciones nuevamente debe ser tomado muy en cuenta pues las combinaciones entre éstas podrían ser significativas para lograr que las mezclas estén dentro de zonas convenientes de resistencia a la tensión.

Capítulo IV. Aplicación del método en un caso real: Selección de mezclas para el P.H. Las Cruces, Nay.





4.5. Análisis

Con las pruebas dispuestas en forma de matrices entradas→salidas (se comparten entradas para la tripleta de parámetros mecánicos que son salidas) se desarrollaron las corridas con las que se definen las reglas de comportamiento en forma de árboles de regresión. La finalidad inmediata es la de encontrar el o los patrones inmersos en la base de datos y conseguir una plataforma para analizar a las variables en actuación simultánea para que, de esta manera, se verifique el impacto de los cambios en su valor sobre el comportamiento de las salidas (variables independientes→variables dependientes VI→VD).

Las reglas de clasificación y/o árboles de decisión-regresión permiten predecir de manera aproximada las salidas propuestas a partir de una combinación determinada. Cabe mencionar que alrededor de los rangos explorados de entradas es que las conclusiones numérico-conceptuales tienen validez. Es posible que alguna entrada pudiera tener una distinta influencia al utilizarse fuera del rango explorado. Sin embargo, las pruebas se movieron en los límites que fueron recomendados por los diseñadores, ingenieros gabinete, especialistas en reconocimiento en campo y consultores de tal forma que son más útiles que si hubieran sido definidos de manera arbitraria o demasiado general.

A continuación se presenta el resultado del análisis por cada salida evaluada.

Resistencia a la compresión en curado acelerado, f_c C/A

- Representación numérica obtenida mediante el algoritmo M5 -árboles de regresión (Quinlan 1992) programado en un software de apoyo.

```
Classifier output
M5 pruned model rules
(using smoothed linear models) :
Number of Rules : 1

Rule: 1

fc C/A MPa =
      -18.6945 * w/c
      + 35.6841 [34/31.094%]
```

Figura 4-4. Regla de clasificación para f_c C/A

Para ésta salida, la resistencia a la compresión en curado acelerado, se encontró que la variable que dirige el comportamiento en mayor medida es la relación w/c . Esto no debe ser traducido como dependencia única pues este resultado depende de que las otras entradas se mantengan estrictamente en los rangos (de preferencia medios) mostrados para cada entrada. La regla de clasificación (en este caso regla de regresión porque la salida es un número y no una clase) se muestra en la *Figura 4-4*. No se obtuvo un tronco con posible toma de decisión (ramas) lo que significa que no se encuentran límites o fronteras de comportamiento, es decir, una expresión simple puede expresar la dependencia paramétrica con suficiente aproximación y la descripción del comportamiento de esta salida puede dirigirse por un mismo camino. Si se usa este modelo lineal para predecir la f_c C/A se logran valores muy cercanos a los medidos en laboratorio.

Se realizó un segundo ejercicio de clasificación sobre esta salida para analizar más profundamente el comportamiento paramétrico de la resistencia la compresión. Cuando se sacrifica el poder de predicción (cálculo de f_c) para maximizar la búsqueda de patrones y la posibilidad de visualizarlos se encuentra un árbol de decisión como el que se muestra en la *Figura 4-5*. Como se observa ahora se pueden declarar que el comportamiento de f_c depende de límites en el valor de w/c , en el nivel más alto en 0.95 y el segundo en 1.2 (mayor cantidad de agua). Una vez pasado el segundo nivel de w/c la variable que marca la diferencia en el valor de f_c C/A de salida depende del tipo de agregado, siendo el agregado CP (triturado) el que dirige los valores máximos de resistencia (mientras las otras variables se mantengan dentro de los rangos medios previamente descritos).

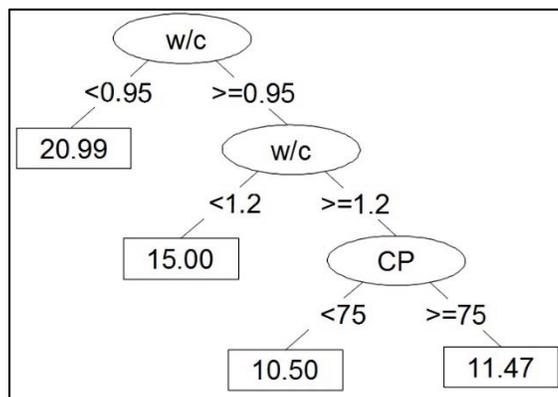


Figura 4-5. Árbol de decisión para f_c C/A

Resistencia a la compresión a 28 días en curado estándar, $f_c 28d$

- Representación numérica obtenida mediante el algoritmo M5 -árboles de regresión (Quinlan 1992) programado en un software de apoyo.

```
M5 pruned model rules
(using smoothed linear models) :
Number of Rules : 1

Rule: 1

fc 28d MPa =
    0.0491 * CPO
    + 0.0338 * CPEG
    - 0.3518 * Agua libre l/m3
    + 0.2023 * Cemento kg
    + 0.0611 * Ceniza volante kg
    + 3.9874 * w/c
    - 0.5827 * Fraccion II
    + 35.1822 [36/37.039%]
```

Figura 4-6. Regla de clasificación para $f_c 28d$

Al igual que en la resistencia a compresión en curado acelerado con un modelo lineal (una regla de clasificación) se logra describir el valor de salida. Lo que es valioso en este ejercicio es la comparación entre este modelo y el obtenido para $f_c C/A$. A pesar de que una revisión superficial (y tradicional) podría llevar a pensar que ambos f_c son completamente relacionables (sólo los separa la escala tiempo), sobre lo expresado en esta base de datos, se concluye que el patrón de comportamiento para $f_c 28d$ depende de más factores entre los que destacan (por su mayor peso) el contenido de cemento, el agua, y el agregado de 9.5-25 mm de tamaño. La ceniza volante y el tipo de cemento tienen impacto sobre el valor de resistencia a la compresión pero lo hacen en un nivel por debajo de los mencionados anteriormente. Utilizar esta ecuación genera valores de f_c muy cercanos a los obtenidos en laboratorio (fase prueba o validación) por lo que se considera oportuna la recomendación de uso.

Para esta salida se realizó también un segundo ejercicio de clasificación para analizar más profundamente el comportamiento paramétrico de la resistencia a la compresión a 28 días. Al sacrificar el poder de predicción para encontrar la dependencia entre variables y patrones, el árbol de decisión que se genera es como el mostrado en la *Figura 4-7*.

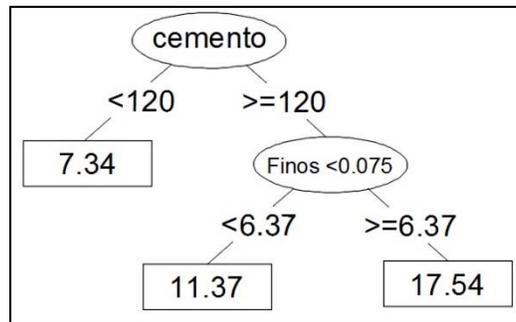


Figura 4-7. Árbol de decisión para f_c 28D
Algoritmo C4.5 (Quinlan, 1993)

En el árbol de la Figura 4-7 se observa que la variable de partida (directora) para predecir la resistencia a compresión a 28 días es el contenido de cemento, con límite de comportamiento en los 120 kg. Interesante es la aparición de los agregados Finos como la segunda variable de control después del contenido de cemento, lo que está de acuerdo con lo que se espera teóricamente y alerta sobre su impacto en el valor final de f_c . La lectura para el ingeniero diseñador de mezclas debe ser entendida como la explotación de estas dos variables para conseguir mejores valores de resistencia mientras las demás permanezcan en los rangos acotados anteriormente.

Módulo de elasticidad en curado acelerado, $E_{40\%}$ C/A

- Representación numérica obtenida mediante el algoritmo M5 -árboles de regresión (Quinlan 1992) programado en un software de apoyo.

Classifier output	Classifier output
Number of Rules : 4	
Rule: 1	Rule: 3
IF	IF
w/c > 0.85	Agua libre l/m3 <= 118.075
Cemento kg <= 100	THEN
THEN	E40 C/A GPa =
E40 C/A GPa =	-0.233 * Agua libre l/m3
-0.012 * CP	+ 46.1694 [8/43.275%]
- 0.0995 * Agua libre l/m3	
+ 0.0355 * Cemento kg	Rule: 4
- 5.3716 * w/c	E40 C/A GPa =
- 0.3842 * Fraccion V	0.4262 * Agua libre l/m3
+ 0.702 * Finos <0.075mm	- 34.7505 [4/35.523%]
+ 31.228 [12/27.261%]	
Rule: 2	Time taken to build model: 0.05 seconds
IF	=== Predictions on training set ===
w/c <= 0.85	
THEN	inst#, actual, predicted, error
E40 C/A GPa =	1 12.9 13.056 0.156
-0.0107 * CP	2 17.2 15.019 -2.181
- 0.1393 * Agua libre l/m3	3 17 16.071 -0.929
- 8.3749 * w/c	4 25.6 22.741 -2.859
+ 45.8571 [10/48.596%]	5 20 22.504 2.504
	6 14.1 15.169 1.069
	7 15.4 15.486 0.086

Figura 4-8. Reglas de clasificación para $E_{40\%}$ C/A

El comportamiento para la salida $E_{40\%}$ C/A es más complejo que para los casos fc. Se genera un árbol de regresión entero que ramifica hasta en 4 modelos lineales de salida. En la *Figura 4-8* se muestran las expresiones (ecuaciones) que permiten calcular el valor del módulo de acuerdo con las condiciones que cumplen sus entradas (ramas).

La representación en árbol se muestra en la *Figura 4-9*. Esta sencilla descripción del comportamiento marca en el nivel más alto de dependencia el valor de w/c con frontera de comportamiento en 0.85. Para las mezclas que poseen un valor de w/c por encima de este límite se cuenta con un segundo parámetro director, el cemento (frontera en 100 kg), a partir del cual se toma al agua como la variable que determina el valor final del módulo. En los casos de w/c mayores a 0.85 se tiene que los agregados (incluidos como vector) determinan en forma importante el resultado. En todos los niveles y para todos los casos el origen del agregado es fundamental y son los triturados los que aparecen con impacto numérico.

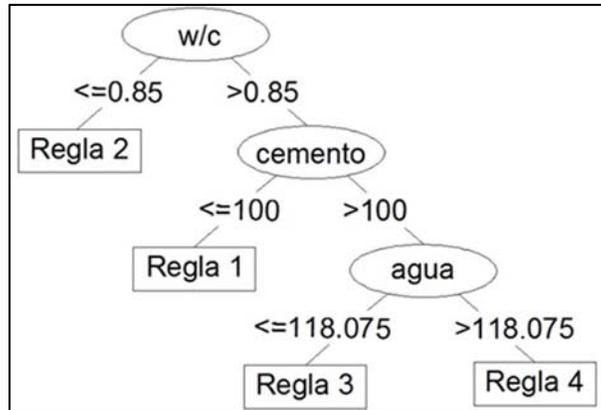


Figura 4-9 Diagrama en base a las reglas de clasificación obtenidas para $E_{40\%}$ C/A

Módulo de elasticidad a 28 días en curado estándar, $E_{40\%}$ 28d

- Representación numérica obtenida mediante el algoritmo M5 -árboles de regresión (Quinlan 1992) programado en un software de apoyo.

```
Number of Rules : 1

Rule: 1

E40 28d GPa =
    0.1198 * Cemento kg
    + 2.3302 * Fraccion III
    - 48.6214 [35/82.439%]
```

Figura 4-10 Regla de clasificación para $E_{40\%}$ a 28d.

Si se comparan los modelos que describen el comportamiento para la salida $E_{40\%}$ C/A y $E_{40\%}$ 28d la complejidad del árbol de regresión con 4 modelos lineales de salida se contrae hasta alcanzar la expresión del módulo con únicamente una regla de clasificación. En esta ecuación son la cantidad de cemento y del porcentaje de Fracción III (2.38-9.5mm) las que marcan en mayor medida el valor de salida.

La expresión numérica de este módulo, cuando está en relación con el total de las otras entradas, no puede ser reducida ni intercambiada cuando se analiza con otros algoritmos de la minería de datos.

Resistencia a la tensión en curado acelerado, f_t C/A

- Representación numérica obtenida mediante software de apoyo.

```
Rule: 1
ft C/A Mpa =
-0.0061 * Ceniza volante kg
- 1.8443 * w/c
+ 3.8481 [35/62.138%]
```

Figura 4-11 Regla de clasificación para f_t en C/A

Para estimar la resistencia a la tensión en curado acelerado la minería de datos genera una regla de clasificación. La expresión del comportamiento de f_t C/A depende de la cantidad de ceniza volante y de la relación agua/contenido cementicio, en mayor medida esta última variable. Importante es resaltar que en esta salida aparece la ceniza como variable a tomar en cuenta si se busca elevar el rango de la f_t .

El árbol de decisión calculado para interpretar este resultado de manera más profunda se muestra en la Figura 4-12.

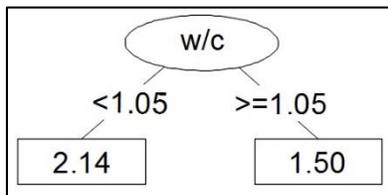


Figura 4-12. Árbol de decisión para f_t en C/A
Algoritmo C4.5 (Quinlan, 1993)

De la misma manera que en los demás esquemas, w/c está presente y como la variable directora. La separación en comportamientos se da en este caso por arriba (ligeramente) de la unidad (1.05). Para las mezclas que están por debajo de esta frontera w/c es el contenido de agregados en Fracción I lo que determina la resistencia a la tensión. El ingeniero diseñador debe tomar en cuenta que, si mantiene los niveles de las otras entradas para conseguir ciertos módulos o valores de f_c , si modifica el valor de la ceniza volante podría conseguir f_t distintos.

4.6. Resultados obtenidos con minería de datos versus Resultados determinados en laboratorio

A continuación se presentan las comparaciones entre lo calculado con minería de datos y los parámetros mecánicos que fueron obtenidos en el laboratorio para la vetaría de mezclas probadas. En este punto se verifica que las expresiones y/o reglas obtenidas del análisis sean congruentes con los resultados obtenidos en laboratorio, o en qué nivel lo son. Para calificar a cada modelo se tomará como apoyo las siguientes graficas:

Resistencia a la compresión en curado acelerado, f_c C/A

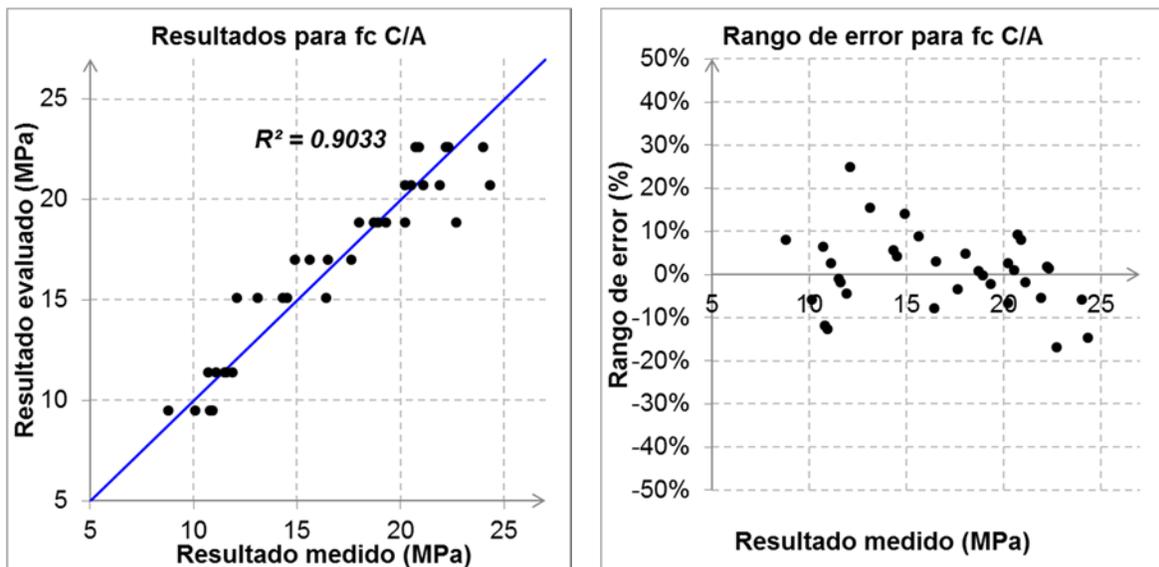


Figura 4-13 Resultados y rango de error para f_c C/A

La relación que se obtiene entre los resultados medidos en laboratorio y los resultados evaluados mediante la *minería de datos*, es una relación lineal, con un coeficiente de regresión $R^2=0.9033$.

Además, la pendiente de la recta de la regresión lineal, corresponde a un ángulo de 42° , muy cercano a 45° , es decir, se tiene una relación muy cercana a 1:1.

También se observa que el rango de error es menor al 10% para la mayoría de los casos, y es menor al 20% para prácticamente para todos los casos, lo que se puede traducir en que se cuenta con una buena predicción de f_c C/A.

Resistencia a la compresión a 28 días en curado estándar, $f_c 28d$

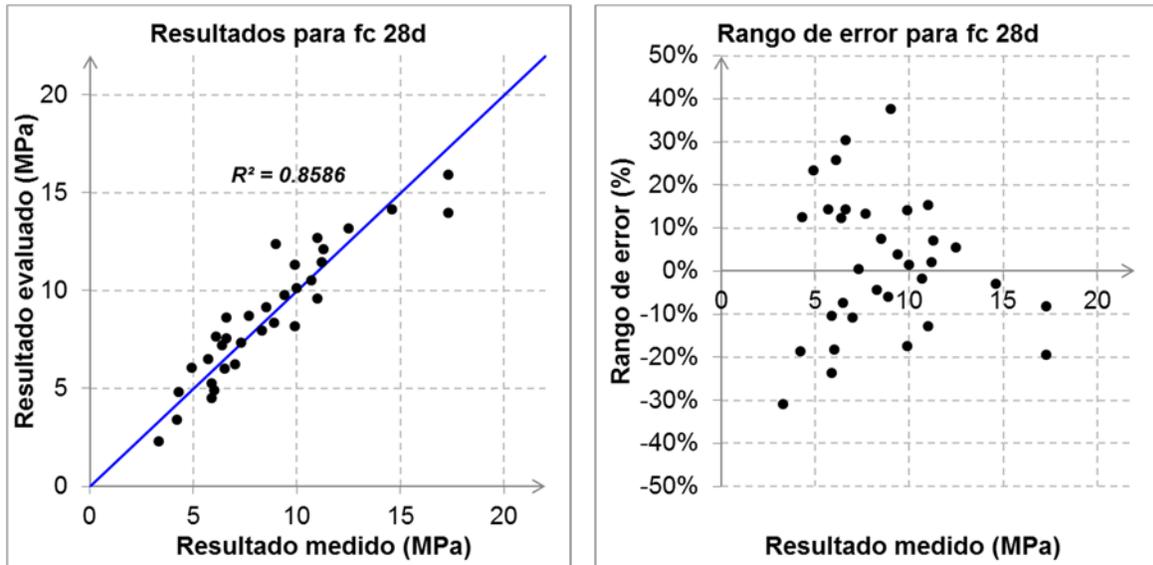


Figura 4-14 Resultados y rango de error para $f_c 28d$

La relación que se obtiene para $f_c 28d$ al igual que $f_c C/A$ es lineal, con un coeficiente de regresión alto ($R^2=0.8586$).

De igual manera, la pendiente de la recta de la regresión lineal, corresponde a un ángulo de 42° , muy cercano a 45° , también se puede decir que se tiene una relación muy cercana a 1:1.

Sin embargo, se tiene mayores rangos de error de los resultados, ya que se cuenta con muchos casos que rebasan el 20% e incluso en algunos casos se alcanza casi el 40% de error respecto al resultado obtenido en laboratorio.

Módulo de elasticidad en curado acelerado, $E_{40\%}$ C/A

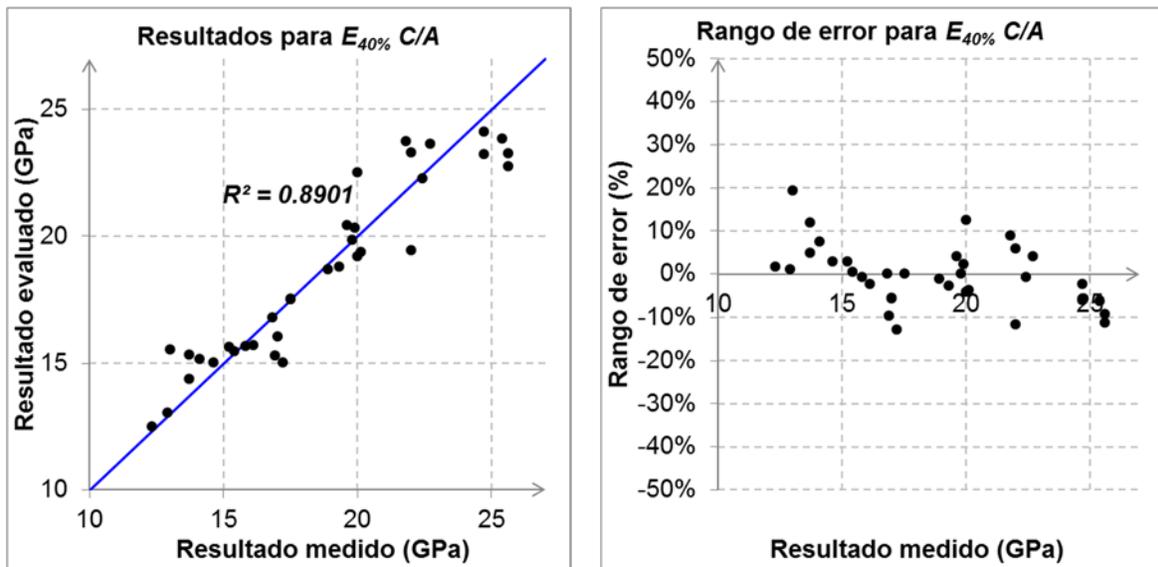


Figura 4-15 Resultados y rango de error para $E_{40\%}$ C/A

La relación que se obtiene para el $E_{40\%}$ C/A también es lineal, con un coeficiente de regresión alto ($R^2=0.8901$).

La pendiente de la recta de la regresión lineal, corresponde a un ángulo de 40° , lo cual corresponde a una relación de 1:0.8, esto significa que el valor evaluado es aproximadamente 20% menor al obtenido en laboratorio, y la predicción ya no es tan exacta, aunque es aceptable.

Se observa que el rango de error es menor al 10% para la mayoría de los casos, y es menor al 20% para todos los casos, lo que se puede traducir en que se cuenta con una buena predicción de la $E_{40\%}$ C/A.

Módulo de elasticidad a 28 días en curado estándar, $E_{40\%} 28d$

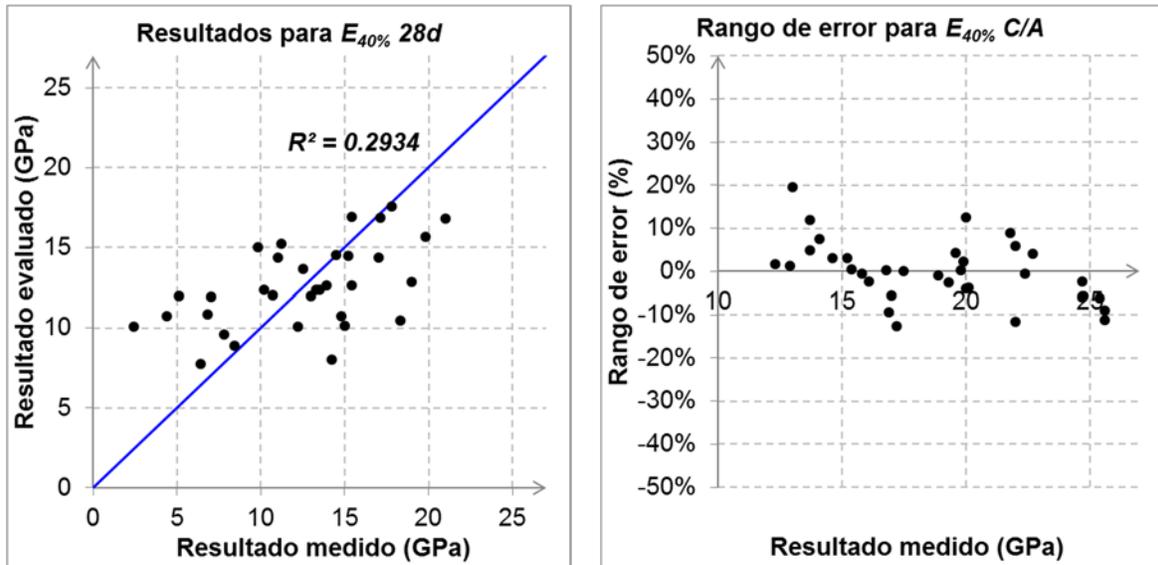


Figura 4-16 Resultados y rango de error para $E_{40\%} 28d$

El coeficiente de regresión obtenido para $E_{40\%} 28d$ es muy bajo, siendo $R^2=0.2934$, se tiene mucha dispersión.

La pendiente de la recta de la regresión lineal, corresponde a un ángulo de 17° , lo cual corresponde a una relación de 1:0.3, el cual no es aceptable.

Los rangos de error obtenidos, son muy altos.

Éste caso particular, se presenta como una oportunidad para repetir el análisis cuando se amplió la base de datos o bien, para otro estudio diferente, ya que los resultados no son aceptables.

Resistencia a la tensión en curado acelerado, f_t C/A

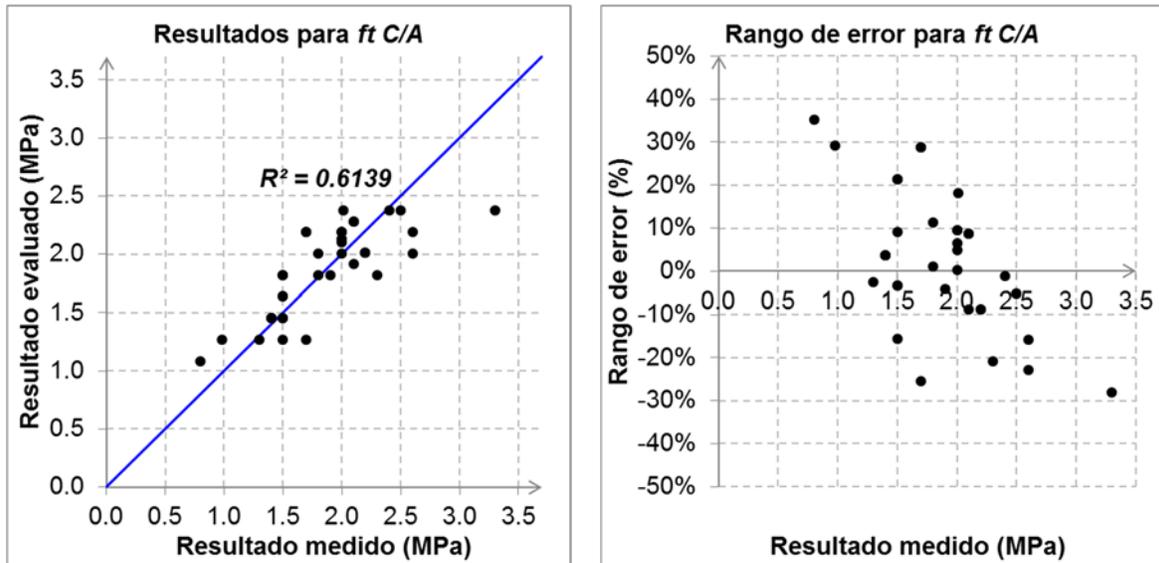


Figura 4-17 Resultados y rango de error para f_t C/A

El coeficiente de regresión lineal que se obtiene para el f_t C/A es aceptable, siendo de $R^2=0.6139$, aunque es bajo.

La pendiente de la recta de la regresión lineal, corresponde a un ángulo de 31° , lo cual corresponde a una relación de 1:0.6, esto significa que el valor evaluado es aproximadamente 40% menor al obtenido en laboratorio, y la predicción ya no es aceptable.

Ésta salida presenta rangos de error altos, teniendo varios casos que superan el 20% e incluso algunos son muy cercanos al 40%.

Los resultados obtenidos para la f_t C/A, si bien no son muy exactos, bien podrían servir de punto de partida y podría retroalimentarse para que una vez que se cuente con una base de datos más extensa, se realice nuevamente el análisis con la intención de obtener resultados más satisfactorios.

4.7. Ejemplos

Para ilustrar el método de utilización de las reglas de clasificación obtenidas, a continuación se presentan dos ejemplos.

Tomando como punto de partida vea *Figura 4-2* y la *Tabla 4-1*:

Tabla 4-1 Parámetros para cada zona de la cortina

Zona	f _c MPa	E Gpa	F _t MPa
A	11.4	13.3	1.4
B	13.4	14.7	1.6
C	18.1	18.2	2.2

Los valores de la *Tabla 4-1* son especificados para un año de edad, sin embargo, se tomarán dichos valores como requerimientos para el Curado Acelerado con la finalidad de ejemplificar el método.

Ejemplo 1.

En este primer ejemplo, se determinarán los componentes requeridos para obtener los parámetros solicitados para la mezcla de la Zona C de acuerdo con la *Tabla 4-1*.

Para la resistencia a compresión en curado acelerado se tiene que:

$$f_c C/A = -18.6945 \cdot w/c + 35.6841$$

Ecuación 4-1 Regla de Clasificación para f_c (C/A)

Si para la Zona C, **f_c C/A = 18.1 MPa**

$$\therefore w/c = \frac{f_c - 35.6841}{-18.6945}$$

$$\therefore w/c = \frac{18.1 - 35.6841}{-18.6945} = 0.94$$

Ahora bien, para la resistencia a la tensión en curado acelerado se tiene:

$$f_t C/A = -0.0061 f.a. - 1.8443 w/c + 3.8481$$

Ecuación 4-2 Regla de Clasificación para f_t (C-A)

Si para la Zona C, **f_t C/A = 2.2 MPa**, y considerando que ya se ha determinado **w/c=0.94**

Se determina la cantidad de ceniza volante (f.a.) como:

$$f.a. = \frac{f_t + 1.8443 \cdot w/c - 3.8481}{-0.0061}$$

$$f.a. = \frac{2.2 + 1.73336 - 3.8481}{-0.0061} = 14 \text{ kg}$$

Por último, para el caso del módulo de elasticidad en curado acelerado la determinación de sus componentes es un poco más complejo ya que se obtuvieron 4 reglas de

clasificación, la aplicación de una u otra depende de ciertas condiciones, por lo cual habrá que observar la *Figura 4-18*:

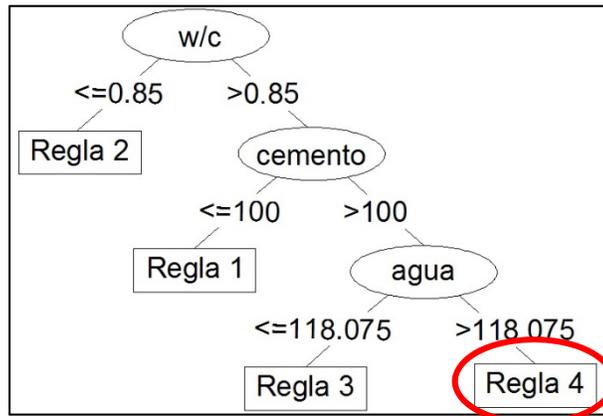


Figura 4-18. Árbol de decisión basado en las reglas de clasificación obtenidas para E_{40%} C/A

a) La primera condición es si **w/c <= 0.85**, o si **w/c > 0.85**.

Se ha determinado que **w/c=0.94**.

b) La segunda condición es si el contenido de cemento "**c**" **<=100 kg/m³**, ó **c>100 kg/m³**.

Para esto se tomará como apoyo la *Figura 4-19*, en la cual se muestran las líneas de tendencia obtenidas del estudio de mezclas de CCR con el cual se elaboró la base de datos. En dicha gráfica se muestran dos tendencias, una corresponde a mezclas de CCR con cemento (línea azul), y la otra línea de tendencia corresponde a mezclas de CCR con cemento+ceniza volante (línea roja).

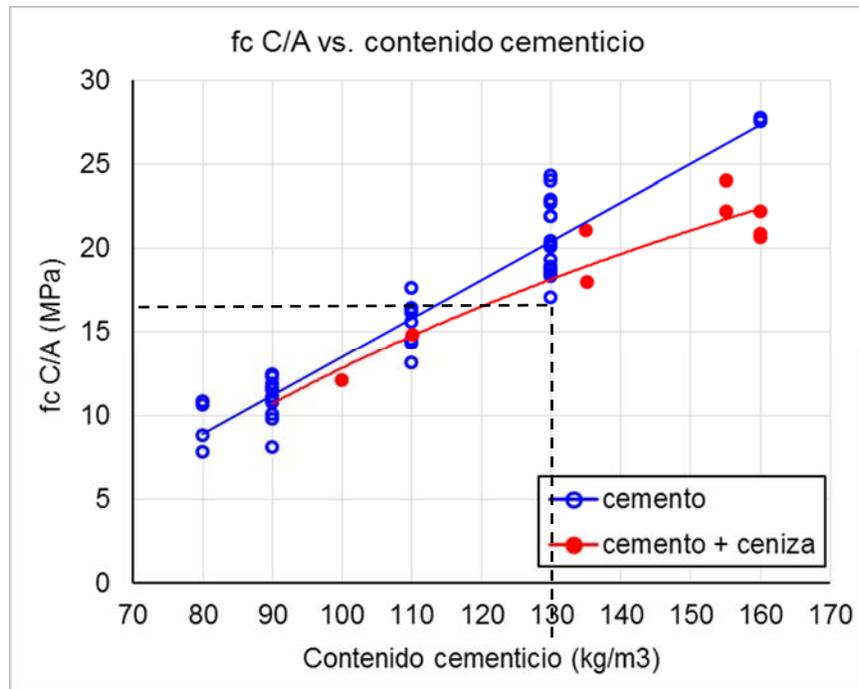


Figura 4-19. Gráfica resistencia a compresión vs contenido cementicio para ejemplo 1
(Obtenida del total de mezclas incluidas en el estudio)

De la Figura 4-19 se deduce que para $f_c=18.1$ MPa (Zona C), el contenido cementicio para mezclas de CCR con cemento + ceniza volante, es de aproximadamente 130 kg/m^3 .

Considerando que $w/c=0.94$, $f.a.= 14 \text{ kg/m}^3$ y contenido cementicio $c.c.=130 \text{ kg/m}^3$, entonces:

$$\text{contenido de cemento } c = 130 - 14 = 116 \text{ kg/m}^3, >100 \text{ kg/m}^3$$

- c) La última condición es conocer si el agua libre $w \leq 118.075 \text{ l/m}^3$, ó $w > 118.075 \text{ l/m}^3$, entonces:

$$w/c.c. = \text{agua libre, } w/c.c.$$

$$0.94 = w/130$$

$$\therefore w = 0.94 \times 130 = 122.2 \text{ lts/m}^3$$

- d) Con estos valores, se decide que se utilizará la regla de clasificación no. 4 de acuerdo con la Figura 4-18

$E_{40\%}CA = 0.4262 \cdot w - 34.7505$ <p style="text-align: center;"><i>Ecuación 4-3 Regla de Clasificación no. 4 para E40% (C/A)</i></p>	<p>Desarrollando la expresión de la Ecuación 4-3, se tiene que</p> $E_{40\%} C/A = 0.4262 \cdot (122.2) - 34.7505$ $\mathbf{E_{40\%} C/A = 17.33 GPa}$
---	--

Se observa que $E_{40\%} C/A$ es ligeramente menor al valor solicitado (18.2 GPa, *Tabla 4-1*), por lo que habrá que hacer un ajuste para alcanzar los requerimientos para la Zona C.

Utilizando la *Tabla 4-1*, y despejando el agua libre, w , se tiene que:

$$w = \frac{E_{40\%}CA + 34.7505}{0.4262} = \frac{18.2 + 34.7505}{0.4262} = 124.24 \text{ lts/m}^3$$

Considerando $w/c = 0.94$, y $w=124.24 \text{ lts/m}^3$, el contenido cementicio y el contenido de cemento serán:

$$\text{contenido cementicio, } c.c. = \frac{124.24}{0.94} = 132 \text{ kg/m}^3$$

$$\text{contenido de cemento, } c = c.c. - fa = 132 - 14 = 118 \text{ kg/m}^3$$

En la *Tabla 4-2* se muestra un resumen con los datos de entrada determinados a partir de las reglas de clasificación, y los resultados esperados:

Tabla 4-2 Resumen del ejemplo 1.

Datos de entrada	Resultados evaluados:	Requerimientos Para Zona C
Cemento c: 118 kg/m³	<i>f_c</i> C/A: 18.1 MPa	<i>f_c</i> C/A: 18.1 MPa
Ceniza volante, <i>f.a.</i> : 14 kg/m ³	<i>f_t</i> C/A: 2.2 MPa	<i>f_t</i> C/A: 2.2 MPa
Contenido cementicio, c.c.: 132 kg/m³	<i>E</i> _{40%} C/A: 18.2 GPa	<i>E</i> _{40%} C/A: 18.2 GPa
Agua w: 124.24 lts/m³		
Relación agua/cont. cementicio w/c.c.: 0.94		

Ejemplo 2.

Se retomará el proceso del ejemplo 1, ahora para el caso de la mezcla de la Zona A según la *Tabla 4-1*.

Nuevamente, para la resistencia a compresión en curado acelerado se tiene que:

$$f_c C/A = -18.6945 \cdot w/c + 35.6841$$

Ecuación 4-4. Regla de Clasificación para f_c (C/A)

Si para la Zona A, **$f_c C/A = 11.4 \text{ MPa}$**

$$\therefore w/c = \frac{f_c - 35.6841}{-18.6945}$$

$$\therefore w/c = \frac{11.4 - 35.6841}{-18.6945} = 1.29$$

Para la resistencia a la tensión en curado acelerado se tiene:

$$f_t C/A = -0.0061 f.a. - 1.8443 w/c + 3.8481$$

Ecuación 4-5 Regla de Clasificación para f_t (C-A)

Si para la Zona A, **$f_t C/A = 1.7 \text{ MPa}$** , y considerando que ya se ha determinado **$w/c=1.29$**

Se determina la cantidad de ceniza volante (*f.a.*) como:

$$f.a. = \frac{f_t + 1.8443 \cdot w/c - 3.8481}{-0.0061}$$

$$f.a. = \frac{1.7 + 2.379 - 3.8481}{-0.0061} = 37 \text{ kg}$$

En este punto se tomará como apoyo la *Figura 4-20*, ya que de los ensayos realizados en laboratorio, se obtuvo que para obtener un **$f_c=11.4 \text{ MPa}$** , se requiere de ~93kg de contenido cementicio considerando la inclusión de ceniza volante a la mezcla, es decir, utilizando la curva roja de la *Figura 4-20*.

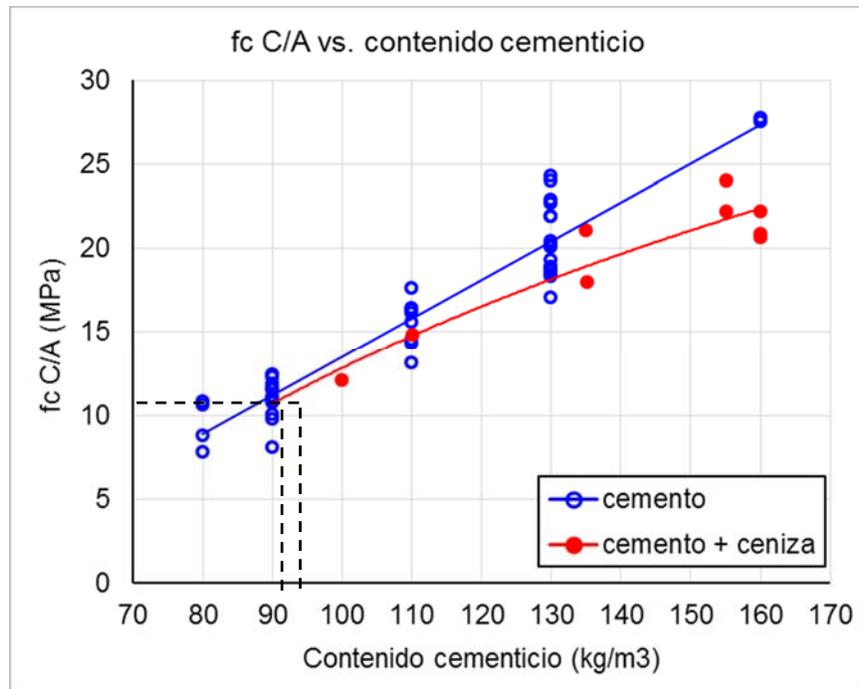


Figura 4-20. Gráfica resistencia a compresión vs contenido cementicio para ejemplo 2 (Obtenida del total de mezclas incluidas en el estudio)

Ahora considerando que en éste ejercicio se requiere de 37 kg/m^3 de ceniza volante, entonces resulta que el contenido de cemento sería de 56 kg/m^3 . El cual está por debajo del contenido estudiado e incluso es un contenido de cemento bajo.

Por lo que se propone agregar el contenido de cementicio totalmente con cemento y así el contenido cementicio sería igual al contenido de cemento, que utilizando la línea de la Figura 4-20 correspondiente a mezclas sin ceniza volante (línea azul) el cual sería de 91 kg/m^3 .

Entonces utilizando la Ecuación 4-5, w/c será:

$$f_t = -1.8443 \cdot w/c + 3.8481$$

$$w/c = \frac{f_t - 3.8481}{-1.8443} = \frac{1.7 - 3.8481}{-1.8443} = 1.16$$

La resistencia a compresión en curado acelerado será de:

$$f_c = -18.6945 \cdot w/c + 35.6841 = -18.6945(1.16) + 35.6841 = 13.99 \text{ MPa}$$

El contenido de agua que se debe agregar a la mezcla es de:

$$\text{agua libre, } w = 1.16 \times 91 = 105.56 \text{ l/m}^3$$

Utilizando la Figura 4-21, considerando $w/c=1.16$, cemento $c=91 \text{ kg/m}^3$, y agua libre $w=105.56 \text{ l/m}^3$, se determina que para obtener el módulo de elasticidad en curado acelerado se debe utilizar la regla de clasificación no. 1.

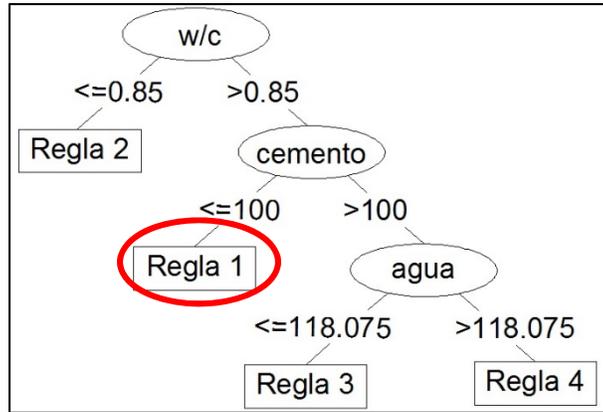


Figura 4-21. Árbol de decisión basado en las reglas de clasificación obtenidas para E40% (C/A)

$$E_{40\%}C/A = -0.012 * CP - 0.0995 * w + 0.0355 * c - 5.3716 * W/c - 0.3842 \text{ Fraccion V} + 0.702 * \text{ finos} + 31.228$$

Ecuación 4-6 Regla de clasificación no. 1 para E_{40%} C/A

De la regla de clasificación mostrada en la Ecuación 4-1, falta por determinar el % de agregado CP, el % de la Fracción V, y el % de finos que se agregarán a la mezcla, los cuales serán:

- Para el % de agregado CP se utilizará la condición más desfavorable, CP=100%.
- Para el caso del % de Fracción V y el % de finos se hará un gráfico con base en los rangos explorados en el estudio. La gráfica se muestra en la Figura 4-22.

Tabla 4-3 E_{40%} en función de los %'s de fracción V y finos

		Finos <0.075 mm, %		
		min	med	max
Fraccion V, %		5.01	6.592	8.2
min	11.48	15.6	16.7	17.9
med	13.014	15.0	16.2	17.3
max	16.19	13.8	14.9	16.1

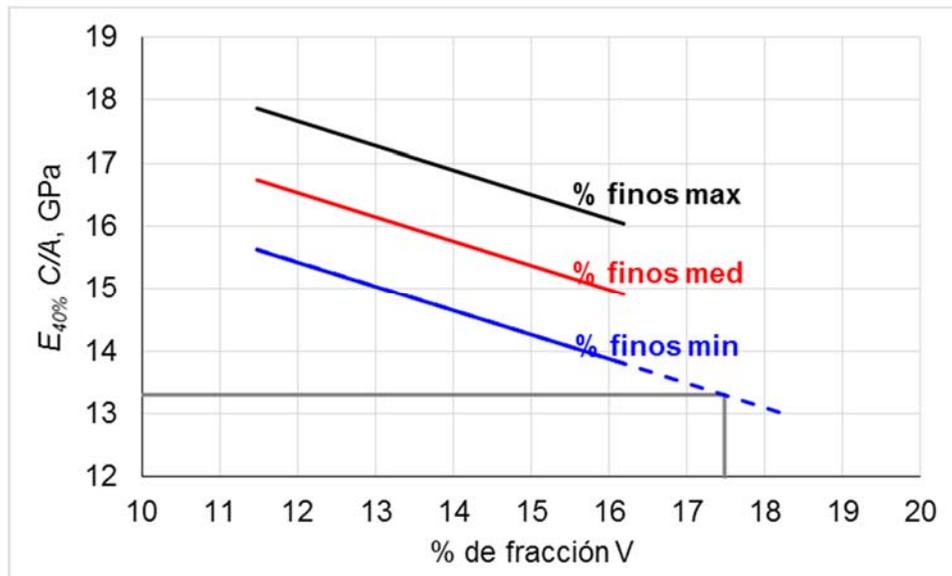


Figura 4-22 $E_{40\%}$ en función del % de la fracción V y finos agregado a la mezcla

Debido a que se conoce el valor de $E_{40\%}$ C/A que se quiere obtener, y al no tener fijos los %'s de fracción V y finos, se puede utilizar la gráfica de la *Figura 4-22* para encontrar la combinación de los %'s de fracción V y finos con la cual se obtiene el valor deseado.

De la *Tabla 4-3* y la *Figura 4-22* se observa que para todas las combinaciones utilizadas el valor obtenido de $E_{40\%}$ C/A, es mayor al deseado, solo una combinación que es utilizar el % máximo de Fracción V con el % mínimo de finos arroja un resultado igual a 13.8 MPa el cual es muy cercano al buscado (13.3 MPa).

Para igualar el resultado obtenido con el resultado buscado, se extrapolará la línea azul correspondiente al % de finos mínimo (aunque se rebasaría el % máximo de Fracción V), para encontrar que para el 17.9% de Fracción V el $E_{40\%}$ C/A = 13.3 MPa. Éste tipo de iteraciones o ajustes quedan sujetos a criterio del diseñador de la mezcla.

En la *Tabla 4-4* se muestra un resumen con los datos de entrada determinados a partir de las reglas de clasificación de $E_{40\%}$ C/A, y los resultados esperados:

Tabla 4-4 Resumen del ejemplo 2.

Datos de entrada	Resultados esperados:	Requerimientos para Zona C
Cemento c: 91 kg/m³	<i>f_c</i> C/A: 13.9 MPa	<i>f_c</i> C/A: 11.4 MPa
Ceniza volante, f.a.: 0 kg/m³	<i>f_t</i> C/A: 1.7 MPa	<i>f_t</i> C/A: 1.7 MPa
Contenido cementicio, c.c.: 91 kg/m³	<i>E</i> _{40%} C/A: Opción 1: 13.8 GPa	<i>E</i> _{40%} C/A: 13.3 GPa
Agua w: 105.56 l/m³		
Relación agua/cont. cem. w/c: 1.16	Opción 2: 13.3 GPa	
% agregado Cerro Partido: 100%		
Opción 1. Fracción V: 16.19% Finos :5.01%		
Opción 2. Fracción V: 17.9% (fuera del rango explorado) Finos :5.01%		

De ambos ejemplos, es importante comentar, que si bien los componentes mencionados en la columna de “datos de entrada” de la *Tabla 4-2* y *Tabla 4-4* son los que más inciden en las salidas evaluadas, también es importante tener en cuenta, que las expresiones empleadas para su cálculo (reglas de clasificación y/o árboles), son parte de la interacción simultánea con el resto de las variables de entrada, es decir, que las reglas de clasificación son válidas para la base de datos utilizada, por lo que el resto de las entradas que no son utilizadas directamente para el cálculo de las “salidas”, deben cumplir con las características y rangos utilizados en la base de datos.

CONCLUSIONES

- Se generaron reglas de clasificación y/o árboles de regresión para identificar los patrones en las mezclas de CCR de los que depende la resistencia a la compresión, el módulo de elasticidad y la resistencia a la tensión indirecta.
- Se encontraron reglas de clasificación y/o árboles de regresión que permiten simular mezclas de CCR y que predicen, de manera aproximada, el comportamiento mecánico de las combinaciones: la metodología usa combinación de entradas (componentes y sus condiciones) para definir parámetros mecánicos de salida. Estas funciones e interpretaciones (bastante simples, *transparentes*) permiten al ingeniero diseñador simular escenarios de manera económica e indagar sobre combinaciones no probadas en el laboratorio.
- De acuerdo con las reglas de clasificación y los árboles de regresión, la relación agua/cemento y el contenido de cemento son las variables que dirigen el comportamiento mecánico de estas mezclas de CCR. En el caso de los módulos elásticos los porcentajes de algunas fracciones de agregado utilizadas también son importantes y para estimar la respuesta a la tensión es importante el valor de la ceniza volante. Por lo que, de acuerdo con la minería de datos, la directriz del proceso o metodología general para monitorear el comportamiento de las mezclas debe ser particularmente enfática en la selección/control cuidadosa(o) de la relación agua/cemento y el contenido de cemento, mientras que se verifiquen los niveles medios de los demás componentes. Si lo que se desea, por algún tipo de restricción, es modificar el contenido de alguna de las variables no directoras, a través de los ambientes de exploración puede simularse el impacto de esta acción sobre las propiedades mecánicas y también se pueden administrar cambios estratégicos manteniendo fijos los niveles de salida (necesidades de diseño).
- Al comparar los resultados evaluados mediante los árboles de regresión con los obtenidos en laboratorio, para la resistencia a compresión en curado acelerado, se obtienen altos coeficientes de correlación (~ 0.90) y para el módulo de elasticidad se tienen rangos de error menores al 10%; la resistencia a compresión a 28d presenta una ligera reducción en correlación pero se mantiene en aproximaciones suficientes para los objetivos del proyecto.

- En el caso de la tensión indirecta, si bien no se obtiene un coeficiente de correlación tan alto como para las salidas descritas en el párrafo anterior, se obtiene una buena predicción del comportamiento, con coeficiente de correlación ~ 0.7 . Este caso es candidato a futura evaluación con redes neuronales. Se intuye que la no linealidad del comportamiento hace que la minería falle en sus predicciones. Las redes permitirán generar mejores aproximaciones.

- Se alerta al lector sobre :
 - i) aunque algunas entradas como el caso de las Fracciones de agregados I, II y IV, y el tipo de agregado SP, no aparecen en las reglas de clasificación ni en los árboles de decisión, esto no significa que no sean necesarias para lograr esos niveles de módulo de elasticidad o de resistencia, sino que los niveles medios en los que fueron probadas son los que desarrollan esos niveles de salida;

 - ii) aunque se hable de respetar los rangos y contener los resultados, la minería de datos tiene poder de extrapolación dentro de ciertos límites, así que se propone al lector interesados en usar con cautela las funciones aquí expuestas para determinar respuestas (parámetros mecánicos) que podrían estar ligeramente fuera de los límites de la base de datos y;

 - iii) el razonamiento expuesto aplica a cualquier obra civil que utilice concreto y en la que se requiera conocer *a priori* las combinaciones que cumplen con el diseño sin inversiones mayúsculas.

- El modelo debe retroalimentarse y probarse contra más casos no-vistos conforme se avance en los estudios previos del PH. La utilización de la minería de datos nos permite tomar mejores decisiones sobre qué aspectos deberían seguir en los ensayos de laboratorio y cuales han sido agotados para los fines del PH.

- Las reglas de clasificación y árboles de regresión son válidos para el PH y los rangos paramétricos enunciados pero el método, los criterios y las experiencias que aquí se presentan puede ser reinterpretadas para otras situaciones de obras civiles en las que se cuente con escasa información y resulte prohibitivo desarrollar numerosas pruebas de laboratorio y/o sondeos exploratorios.

REFERENCIAS

American Concrete Institute (ACI) 1999. “ACI 207.5R99 Roller-Compacted Mass Concrete”. American Concrete Institute, 1999

American Society for Testing and Materials (ASTM) 1998. “C1170 Standard Test Methods for Determining Consistency and Density of Roller-Compacted Concrete Using a Vibrating Table”. ASTM 1998.

Breiman, L., Friedman, J., Olshen, R., and Stone, C., 1984. “Classification and Regression Trees”, Monterey, CA: Wadsworth, 1984.

Brent, R., 1990. “Fast Training Algorithms for Multi-Layer Neural Nets” Numerical Analysis Project Manuscript NA-90-03, Computer Science Department, Stanford University, Stanford, CA 94305, 1990.

CFE/GEIC, 2012. “Informe Geológico de avance a nivel de factibilidad de la alternativa Rancho Viejo P.H. Las Cruces, Nayarit”.

Chou, S., et al 2011. “Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques”. Journal of computing in civil engineering. 25: 242-253, 2011.

Escalaya, M., Alva, J. 2001. “Diseño de Mezclas de Concreto Compactado con Rodillo utilizando conceptos de Compactación de Suelos”. XIV Congreso Nacional de Ingeniería Civil 2001.

(Fu, 1994). Fu, L., “Neural Networks in Artificial Intelligence”, New York: McGraw-Hill, 1994.

Garduño, A. 2015. “Informe 15-006-SGM-M Estudio de Mezclas de CCR”. Comisión Federal de Electricidad - Gerencia de Estudios de Ingeniería Civil. 2015

Haykin, S. 2009. “Neural Networks and Learning Machines”. 3a. Ed. Prentice Hall 2009.

Hertz, J., Krogh, A., y Palmer, R., 1991, “Introduction to the Theory of neural Computation”, Lecture Notes vol 1. Santa Fe Institute Studies in the Sciences of Complexity. Addison Wesley, 1991.

Hollingworth, F., Druyts, F.H.W.M. 1986. “Rollcrete: some applications to dams in South Africa Water Power and Dam Construction”. Londres, 1986

International Commission on Large Dams (ICOLD) 2003. “*Bulletín 126 Roller-Compacted Concrete Dams, State of the art and case histories*”. ICOLD 2003.

John, G., Lauderdale, 1995. “*Robust Linear Discriminant Trees*” Proc. Of the Conf. on Artificial Intelligence and Statistics, Ft. Lauderdale, Florida, 1995.

Kohavi, R., 1994. “*Bottom-Up Induction of Oblivious Read-Once Decision Graphs*”. Proc. of European Conference on Machine Learning (ECML-94), 1994.

Langley, P., 1996. “*Elements of Machine Learning*”, San Francisco: Morgan Kaufmann, 1996.

Li, 2015. “*LI CRS 4212 DES OC MC 014, Memoria de cálculo. PH Las Cruces. Obra de Contención. Análisis de Esfuerzos de la Presa*”. Proyecto ejecutivo PH Las Cruces, Nayarit. CFE 2015

Liao, S., et al. 2012. “*Data mining techniques and applications – a decade review from 2000 to 2011*”. Expert Systems with Applications. 39: 11303-11

Maimon, O, Rokach, L. 2010. “*Data Mining and Knowledge Discovery*”. Second Edition. Springer. 2010.

Marchand, M., y Golea, M., 1993. “*On Learning Simple Neural Concepts: From Halfspace Intersections to Neural Decision Lists*”, Network, 4:67-85, 1993.

Murillo, R. 2012. “*Presas de tierra y sus fallas*”. XXVI Reunión Nacional de Mecánica de Suelos e Ingeniería Geotécnica. SMIG. 2012.

Nagayama, I 1991. “*A study on mixture design of RCD by large-sized specimen compaction device - Engineering for Dams, N°62*”. 1991

Natarajan, B., 1991. “*Machine Learning: A Theoretical Approach*”, San Francisco: Morgan Kaufmann, 1991.

Nawy, E. et al 2008. “*Concrete Construction Engineering Handbook*”. 2nd edición. CRC Press Taylor & Francis Group 2008.

Nikoo, M. et al 2015. “*Prediction of concrete compressive strength by evolutionary artificial neural networks*”. Advances in Materials Science and Engineering. ID 849126 2015

Nilsson, N. J. 1998. “*Introduction to machine learning*”. Stanford University 1998.

Oliver, J., Dowe, D., y Wallace, C., 1992. “*Inferring Decision Graphs using the Minimum Message Length Principle*”. Proc. 1992 Australian Artificial Intelligence Conference, 1992.

Pérez, G, Molina, J. 2013. “*Apuntes de Obras Hidráulicas*”. Universidad Michoacana de San Nicolás Hidalgo, Facultad de Ingeniería Civil, Departamento de Hidráulica, 2013.

Quinlan, J. y Riverest, R, 1989. *"Inferring Decision Trees Using the Minimum Description Length Principle"* Information and Computation, 80:227-248, Marzo, 1989.

Quinlan, J. 1986. *"Induction of Decision Trees"*. Machine Learning, 1:81-106, 1986. Reimpreso en Shavlik, J. y Dietterich, T., *"Readings in Machine Learning"*, San Francisco: Morgan Kaufmann, 1990, pp. 57-69.

Quinlan, J. 1987. *"Generating Production Rules from Decision Trees"* In IJCAI-87: Proceedings of the Tenth Intl. Joint Conf. on Artificial Intelligence, pp. 304-7, San Francisco: Morgan-Kaufmann, 1987.

Quinlan, J. 1992. *"Learning with continuous classes"*. 5th Australian Joint Conference on Artificial Intelligence. 343-348. Singapur 1992.

Quinlan, J. 1993. *"C4.5: Programs for Machine Learning"*. San Francisco. Morgan Kaufmann, 1993.

Rissanen, J., 1978. *"Modeling by Shortest Data Description"* Automatica, 14:465-471, 1978.

Schrader, E, Stefanakos, J. 1995. *"RCC cavitation and erosion resistance"*. Proceedings, International Symposium on Roller Compacted Dams, October 2-4, Santander, Spain. pp. 1175-1188.

U.S. Army Corps of Engineers (USACE) 1995. *"Gravity dams design, EM 1110-2-2200"*. USACE 1995.

U.S. Army Corps of Engineers (USACE) 2000. *"Roller-Compacted Concrete, EM 1110-2-2006"*. USACE 2000.

Valencia, E. 2006. *"Diseño de presas de concreto"*. UNAM Tesis de licenciatura 2006.

Wang, X., Luo, S., Yuan, Q., Wang, H. 2012. *"High-Speed flow erosion on a new Roller Compacted Concrete Dam During Construction"*. Journal of Hydrodynamics 2012, 24(1):32-38.

Weiss, S., y Kulikowski, C., 1991. *"Computer Systems that Learn"*, San Francisco. Morgan Kaufmann, 1991.

Witten, I., y Frank, E., 2005. *"Data Mining Practical Machine Learning Tools and Techniques"*. Second Edition. Ed. Elsevier. 2005.

Yeh, I. 2007. *"Computer-aided design for optimum concrete mixtures"* Cement and Concrete Composites. 29: 193-202. 2007.