



UNIVERSIDAD AUTÓNOMA CHAPINGO

DIVISIÓN DE CIENCIAS FORESTALES

DEPARTAMENTO DE ESTADÍSTICA, MATEMÁTICA Y CÓMPUTO

CV Ejecutivo

Contenido

1. Resumen sintético	3
2. Breve reseña histórica de la UACH	4
3. Experiencia y capacidad técnica.....	6
4. Listado de proyectos de investigación.....	7
5. Listado de proyectos de consultoría	8
6. Infraestructura.....	8

1. Resumen sintético

Vinculación con otras instancias

El equipo de investigación que se presenta cuenta ya con poco más de 20 años de experiencia en la investigación y servicio, contribuyendo a través del estudio de grandes volúmenes de datos, al análisis de distintos procesos sociales que ocurren en México y que son del interés público. Esta experiencia ha sido acumulada a través de vinculaciones, principalmente con el Consejo Nacional de Ciencia y Tecnología (CONACyT) y los Fondos Sectoriales de Investigación SEP/SEB, INEGI, INMUJERES y CONAFOR; también con el Gobierno del Estado de Guanajuato; así como instancias de carácter Federal como el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL); y recientemente con el Servicio de Administración Tributaria (SAT). Los resultados del Megaproyecto de Investigación en Estadística Social se difunden en el sitio oficial: <http://demyc.chapingo.mx/idh/>

RRHH e Infraestructura

Todo el trabajo desarrollado y la vinculación realizada ha sido posible gracias a la entusiasta colaboración de jóvenes pasantes o egresados de la Licenciatura en Estadística de la UACH; por lo que su participación en nuestras investigaciones y proyectos es fundamental. Sobre la infraestructura, el DEMyC dispone de un clúster para el manejo de Big Data que ha permitido el uso de la estadística y la ciencia de datos para los estudios que en él se realizan.

2. Breve reseña histórica de la UACH

La Universidad Autónoma Chapingo es una institución mexicana pública, con liderazgo y reconocimiento nacional e internacional con alta calidad académica en la educación; los servicios y la transferencia de las innovaciones científicas y tecnológicas que realiza; la importancia y magnitud de sus contribuciones en investigación científica y tecnológica; y por el rescate y la difusión cultural que desarrolla. Para conocer la oferta de servicios académicos de la UACH, puede consultar su dirección electrónica en: <http://web.chapingo.mx/>

La UACH tiene sus orígenes en la Escuela Nacional de Agricultura (ENA), la cual fue fundada en forma oficial el 22 de febrero de 1854, en el Convento de San Jacinto, D.F. Posteriormente, la ENA se trasladó a la ex hacienda de Chapingo, donde inicia sus actividades el día 20 de noviembre de 1923. En este proceso de cambio, la escuela adopta el lema “ENSEÑAR LA EXPLOTACION DE LA TIERRA, NO LA DEL HOMBRE”, mismo que se encuentra plasmado en su Acta de Inauguración en Chapingo.

Se realizaron cambios muy importantes en 1941, con relación a la política de admisión de la Escuela, pues para ingresar a la misma, se exige como requisito la culminación de los estudios a nivel secundaria; de esta forma, se estableció la Preparatoria Agrícola de tres años y la Especialidad de cuatro años.

El 22 de febrero de 1959, se crea el Colegio de Posgraduados de la Escuela Nacional de Agricultura (más tarde, éste se establece como organismo independiente), lo que constituyó un impulso fundamental para el desarrollo de la educación agrícola superior del país.

Durante los años de 1962-63, la entonces Secretaría de Agricultura y Ganadería, y el director de la ENA decretaron la desaparición de la Preparatoria Agrícola. Sin embargo, en 1966 el H. Consejo Directivo decide reabrirla, favoreciendo con ello el ingreso de alumnos provenientes de las áreas rurales del país; a partir de entonces se acepta el ingreso de estudiantes que han concluido el nivel secundaria o preparatoria.

En 1969, bajo un ambiente democrático, se inicia la desaparición de la disciplina militar, y en 1973 se logra el establecimiento de la autodisciplina. En esta época se gestan importantes reformas en las políticas del gobierno de la escuela, y son presentados dos proyectos para la transformación de la escuela a universidad; éstos fueron el UNRAM y el UNACH, siendo el último el que fue enviado a las Cámaras Legislativas del país para su análisis y aprobación. El 30 de

diciembre de 1974, aparece en el Diario Oficial de la Federación, la promulgación de la “LEY QUE CREA LA UNIVERSIDAD AUTONOMA CHAPINGO”.

El proceso de transformación de escuela a universidad culmina en 1978, con la formulación del Estatuto de la Universidad Autónoma Chapingo; a partir de entonces todas las actividades y planteamientos políticos, inherentes a la institución, se rigen a través de dicho Estatuto. En este año, inicia sus actividades con los siguientes departamentos: Preparatoria Agrícola, Bosques, Economía Agrícola, Fitotecnia, Industrias Agrícolas, Irrigación, Parasitología Agrícola, Sociología Rural, Suelos, Zonas Áridas y Zootecnia.

Desde su creación, la universidad ha experimentado un proceso de expansión en cuanto a las diferentes orientaciones de la agronomía, tanto a nivel licenciatura como en posgrado. Es así como se han creado las carreras de: Agroecología, Mecánica Agrícola, **Estadística**, Forestal Industrial, Forestal, Restauración Forestal, Administración de Empresas Agropecuarias, Comercio Internacional de Productos Agropecuarios, Economía Agrícola, Planeación y Manejo de los Recursos Naturales Renovables, Agrónomo Especialista en Zonas Tropicales, Sistemas Pecuarios y Sistemas Agrícolas de Zonas Aridas. Dentro del propio seno universitario se han creado las Maestrías en Ciencias en Economía del Desarrollo Rural, Sociología Rural, Producción Animal, Protección Vegetal, Ciencias Forestales, Desarrollo Rural Regional, Horticultura y los Doctorados en Ciencias en Economía Agrícola, y en Ciencias Agrarias.

3. Experiencia y capacidad técnica

La Universidad Autónoma Chapingo, en particular, el Departamento de Estadística, Matemática y Cómputo, en su trayectoria de investigación y servicio cuenta con 20 años de experiencia en la aplicación de estadística y matemática para el manejo y análisis de grandes bases de datos; en particular de lo que ahora se conoce como Big Data. Por citar tres ejemplos:

1) Los resultados del **Informe Nacional de Desarrollo Humano** referidos a 2015, derivan a su vez del procesamiento y análisis de los microdatos de la Encuesta Intercensal 2015, cuyas bases de datos de vivienda contienen 5.8 millones de registros y las relativas a personas poco más de 22.6 millones; entre ambas bases de datos se manejan poco más de 150 variables. De la misma manera, los informes de desarrollo humano previos a 2015 (1995-2010) requirieron en su momento el procesamiento de los datos provenientes de los censos, conteos y muestras de población y vivienda, cuyas bases de datos también pueden clasificarse en el ámbito del Big Data.

2) También se cuenta con experiencia en la conformación de grandes bases de datos, ejemplo de ello son las entregadas al Fondo Sectorial CONACyT-INEGI, con poco más de 1.2 millones de registros y poco más de 100 variables, mismas que derivaron del proyecto de investigación denominado Muestreo Probabilístico para la Recuperación de Microdatos del Censo General de 1930.

3) El estudio realizado para el Servicio de Administración Tributaria en el año 2018, denominado Evasión en Sueldos y Salarios; en el desarrollo de esta investigación se procesó una volumetría de poco más de 30 millones de registros que constituyen la base de datos de la Declaración Informativa Múltiple 2016 y 980 millones de registros en la base de datos del CFDI de nómina.

Cabe aclarar que el procesamiento de información involucrado en los trabajos de investigación y resumido en líneas anteriores, sólo ha sido posible gracias a que la Universidad autónoma Chapingo posee la Infraestructura adecuada y suficiente en sus laboratorios de cómputo; las licencias de software especializado, particularmente de SAS, SPSS, y el software libre R; y recursos humanos altamente especializados.

4. Listado de proyectos de investigación

Entre los proyectos de investigación desarrollados del año 2000 a la fecha están:

- **Permanente. Informe Nacional sobre Desarrollo Humano en México.**
Con esta investigación se busca contribuir al análisis de la dinámica del desarrollo humano en el periodo 1995-2000-2005-2010 -2015 siguiendo las ideas primarias de la Organización de las Naciones Unidas. Los resultados se encuentran disponibles, según la información lo permite, a escala Nacional, Estatal, Municipal y por localidades. Los resultados y publicaciones del informe están disponibles en: <http://demyc.chapingo.mx/idh/>
- **Desarrollo y Calidad Educativa según Grado de Ruralidad en los Municipios de México: 2005.**
Financiado por el Fondo Sectorial de Investigación SEP/CONACyT.
- **Evaluación de la Eficiencia y Eficacia en la Planeación, Seguimiento y Asignación Presupuestal de Programas para la Equidad de Género.**
Apoyado por el Fondo Sectorial de Investigación INMUJERES-CONACyT.
- **Muestreo Probabilístico para la Recuperación del Censo de Población 1930.**
Solicitado en convocatoria por el Fondo Sectorial de Investigación CONACyT-INEGI.
- **Desarrollo de un Sistema Biométrico para la Planeación del Manejo Forestal de los Ecosistemas con Potencial Maderable en México.**
Financiado mediante un consorcio de instituciones por el Fondo Sectorial para la Investigación, el Desarrollo y la Innovación Tecnológica Forestal (CONAFOR-CONACYT). El objetivo principal del proyecto fue generar las ecuaciones alométricas para especies de importancia maderable a nivel de UMAFOR en cada uno de los estados comprendidos en la Estrategia Nacional de Manejo Forestal.
- **Estadísticas sociales para Estudios de Política Pública y de Prospectiva.**
Concurrió y obtuvo financiamiento en la Convocatoria de Investigación Científica Básica del CONACyT.
- **Sistema de Indicadores de Vivienda.**
Solicitado en convocatoria por el Fondo Sectorial de Investigación CONACyT-INEGI.
- **Modelos de cambios para estimar la pobreza a nivel municipal a partir de información 2010 y 2015.**
Financiado por el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL)

5. Listado de proyectos de consultoría

Para la Comisión Nacional Forestal (<https://snigf.cnf.gob.mx/encuesta-nacional-de-beneficiarios/>):

- Encuesta Nacional de Beneficiarios de PROÁRBOL 2011
- Monitoreo y Evaluación Complementaria del Programa U036 PRONAFOR-Desarrollo Forestal 2013
- Encuesta Nacional de Personas Beneficiarias de la CONAFOR 2016

Para el Servicio de Administración Tributaria (http://omawww.sat.gob.mx/cifras_sat/Paginas/inicio.html):

Documentos de referencia, Estudios de evasión y elusión

- 2018 Evasión del ISR en sueldos y salarios
- 2019 Prácticas de evasión del ISR en sueldos y salarios
- 2020 Evasión en sueldos y salarios: redes de patrones

6. Infraestructura

La UACH, es la más antigua del país en su especialidad y, habiéndose desarrollado la estadística moderna en el ámbito de la agricultura, posee una fuerte y excelente escuela en esta disciplina; cuenta con laboratorios de cómputo (hardware y software), auditorios, aulas, acervos bibliográficos, tanto impresos como digitales, servicios y productos de información tecnológica, etcétera. Mismos que podrán ser utilizados para el desarrollo de los trabajos relacionados al proyecto.

Por su parte, el DEMyC cuenta con un laboratorio de cómputo, el cual ha permitido llevar a cabo actividades de investigación, permitiendo instrumentar complejas rutinas computacionales, y desarrollar métodos analíticos en el contexto de la ciencia de datos para la resolución de problemas. Para esto, el DEMyC cuenta con un cluster para Big Data constituido por 7 máquinas (1 nodo maestro y 6 nodos esclavos) con una conexión a una red local y transmisión de hasta 1000 Mb/segundo.

Características físicas

La capacidad de procesamiento del cluster para Big Data es de 63 GB de RAM disponibles y 42 núcleos libres para cualquier proceso, los procesadores son Intel core i7 y tienen una velocidad de 3.5 Ghz, cada una. De ser necesario, para las tareas que demanden una mayor capacidad de

procesamiento, el número de nodos esclavos podrá aumentarse para mejorar las características del cluster.

El nodo maestro tiene un disco duro sólido (SDD) de capacidad de 500 Gb.

Los nodos esclavos tienen un disco rígido (HDD) con capacidad de 1TB y 500 Gb dedicados.

Características software

Cada nodo tiene un sistema operativo Linux (CentOs 7) en el que se tiene instalado el software Hadoop y Spark, esto último permite el manejo de las bases de datos a través del sistema de archivos compartido de Hadoop (HDFS) y la instalación, de ser necesario, de Hive, Sqoop, Hbase, entre otras.

Para dimensionar la capacidad del cluster, es necesario mencionar que en la página oficial de Apache Hadoop¹ se encuentra un listado de las instituciones y empresas que usan la tecnología de Apache para el análisis del Big Data, en dicha lista se pueden encontrar diversos tipos de *cluster*, algunos de ellos con características de hardware menores a las características con las que se cuentan en el *cluster* asignado para el procesamiento requerido. Algunos de estos *cluster* tienen 4 nodos, RAM de 4 a 6 GB por nodo y un almacenaje total de 1 a 4Tb.

En caso de ser requerido para las tareas de procesamiento, el número de nodos esclavos podrán aumentarse para mejorar las características del *cluster*.

Asimismo, el DEMyC cuenta con un sistema de computadoras para la instrumentación de High Performance Computing (HPC) cuyo propósito es el procesamiento en paralelo, a diferencia del cluster descrito anteriormente, el cual está enfocado al procesamiento de grandes bases de datos. Este sistema de computadoras tiene instalado el software R, Python y compiladores de C, C++ y Fortran.

¹ <https://wiki.apache.org/hadoop/PoweredBy>