IMPLEMENTACIÓN DE UN SISTEMA DE MANEJO CLIMATICAMENTE INTELIGENTE (Climate Smart Managment Sistem CSMS), BASADO EN BIG-DATA Y AGRICULTURA ESPECÍFICA POR SITIO (AEPS).

Protocolo estandarizado de captura de información necesaria para CSMS.

Colecta, consolidación y almacenamiento de información útil en	Protocolo No. 0XX
agricultura	
BIG – DATA / AEPS – CIAT	Versión: Julio de 2016
Por: Hugo A. Dorado.	Revisado: DJ / AM

Introducción.

Un productor que haya presenciado una buena cantidad de ciclos de siembra en su finca, puede darse cuenta de la gran variación que experimenta el rendimiento de su cultivo, esto se debe a que la respuesta depende tanto de factores ambientales como del manejo agronómico y ambos cambian frecuentemente en el tiempo.

Ocasionalmente puede incluso tomar decisiones erradas bajo unas condiciones específicas, y en el futuro realizar ajustes para otros ciclos de siembra conforme a la experiencia que va acumulado; del mismo modo puede replicar prácticas con las que haya tenido buenos resultados y así irse orientando a tomar cada vez mejores decisiones.

Este proceso de acumular experiencia puede refinarse un poco más al tratar de medir y registrar información más precisa del cultivo, puede ser que el productor adopte hábitos sencillos desde apuntar las fechas de siembra, cosecha, aplicaciones, monitoreo y georreferenciación; hasta sistematizar los análisis de suelos e instalar estaciones meteorológicas. Muchos ya han comenzado a colectar datos sin embargo en forma separada o desorganizada y se están perdiendo la oportunidad de sacarles provecho, otros aún no han comenzado a registrar información ya sea porque desconocen la manera en que podrían hacerlo o no ven el beneficio a largo plazo que les traería.

En este documento se propone un protocolo para orientar al usuario en como: medir, consolidar y almacenar información útil en agricultura, ya sea en lotes comerciales o experimentales. Está dirigido a técnicos en agricultura, productores o personal de campo que requieran o deseen capturar información de cultivos, para luego ser analizada y sirva como insumo para tomar decisiones en futuro.

Estado del arte de los sistemas productivos.

Al pretender utilizar grandes bases de datos (Big-Data) y la herramienta de Agricultura Específica por Sitio (AEPS) es necesario tener una muy buena percepción del sistema productivo, en términos de su problemática y potencialidades.

Es conveniente tener claros los objetivos para los que se quiere aplicar este enfoque, algunos podrían ser: aumentar productividad, optimizar calidad, identificar prácticas relevantes, zonificar la producción, entre otros. Puesto que de esto depende la manera en que se orienta la estrategia de colecta de variables y métodos de análisis que se van a emplear posteriormente.

Si bien los análisis ofrecerán información, es necesario contar con expertos que validen y puedan dar un contexto coherente, a dicha información, con base en conocimientos agronómicos, fisiológicos o de otra índole en torno al sistema productivo.

La escala o cobertura siempre será un aspecto a tener en cuenta, es importante poder considerar la mayor variabilidad posible de condiciones de producción. Variabilidad en términos de clima, suelo y prácticas de manejo muy asociado a perfiles de productores. Por otro lado, poca variabilidad en factores que explican el rendimiento puede ocasionar que sea más difícil distinguir las características relevantes, dado que no cambia mucho la respuesta y los valores medidos en otras variables son muy homogéneos para los eventos, y como consecuencia tendrán muy pocos aportes en la explicación de la variable de interés.

Identificar tipos y fuentes de datos

AEPS, básicamente parte de integrar información de suelo, clima y prácticas de manejo del cultivo (variables de entrada), para explicar la variabilidad en producción, calidad u otra variable de salida, que sea de interés.

Identificar y reconocer información útil

Para comenzar se debe identificar las variables que son de interés principal y necesarias de acuerdo al objetivo planteado; además que estén involucradas dentro del ciclo productivo del cultivo, por tanto es ideal que el usuario haga una lista en la que se enumeren las variables que debe medir, que ya posee o que debe conseguir por otras fuentes.

En este paso es útil consultar expertos, revisar material bibliográfico y/o recibir retroalimentación con análisis previos. Sin embargo ya se pueden definir unas variables que por experiencia son importantes y es preferible que se incluyan en la lista.

Se tiene que para el caso de *suelos* las variables pueden ser: propiedades químicas (algunos elementos como: N, P o K,...), propiedades físicas (estructura, texturas ó resistencia,...) y/o propiedades funcionales (profundidad efectiva, drenajes,...); en *clima*, las básicas son: temperaturas, precipitación, radiación solar y humedad relativa, pero también pueden ser útiles: velocidad del viento, evapotranspiración, entre otras; en *manejo* generalmente hay numerosas posibilidades, siendo las más relevantes: fecha de siembra, cosecha, variedad, cantidad y fechas de fertilización química, fertilizaciones orgánicas, controles y sistemas de siembra; por último la variable de *salida* que es de principal interés: rendimiento, calidad de cosecha o utilidad económica; en el Anexo 1 de este protocolo se encuentra un diccionario de datos en el que se puede observar una lista de variables sugeridas que han sido utilizadas en arroz y puede servir de ejemplo.

Información necesaria para comenzar

AEPS también puede generar resultados con parte o la totalidad de los componentes de información, lo que supone un proceso continuo de colecta de datos y mejora de los resultados obtenidos; de hecho se puede hacer un primer acercamiento con tan solo datos de siembra, cosecha y estaciones meteorológicas próximas y no es necesario para comenzar tener ya consolidada la información completa que se sugiere en el diccionario de datos.

Explorar fuentes de información secundaria

Información secundaria se refiere al tipo de información que ya fue o viene siendo colectada para otros propósitos y que puede ser útil para hacer AEPS, con ella se puede obtener una aproximación al objetivo deseado, y también permitirá reconocer cuales son las variables faltantes, en este caso sí existen registros en físico o plantillas de Excel, con datos de siembra, aplicación de insumos o cantidad cosechada puede ser una buena base para comenzar. También existe el tipo de información secundaria que es de acceso libre y que la generan entidades de investigación, como lo son los datos de clima y suelo.

Para obtener información de clima, existen bases globales que permitirán descargar datos con cierto nivel de resolución, algunas de ellas: World-Clim (http://www.worldclim.org/), la cual es una base de datos libre con 1 kilómetro de resolución espacial y provee información de temperatura y precipitación, sin embargo no alcanza una resolución temporal ya que es un promedio de 50 años; otra posibilidad es CRU, la cual es una base de datos también de acceso libre (https://crudata.uea.ac.uk/cru/data/hrg/#current), con diversa información climática y resolución de 40 km, pero con la ventaja de que es mensual. Una de las mejores opciones también es contactar la entidad encargada de servicios meteorológicos del país correspondiente, la cual normalmente posee información diaria y una cantidad mayor de variables climáticas, en el caso de Colombia sería IDEAM (http://www.ideam.gov.co/). En la web hay mucha información disponible de clima, alguna paga, otra gratuita, el caso es detectar la información relevante con base en la geolocalización de los lotes y los periodos de interés (Fechas de siembra y cosecha).

Para obtener información de suelos, también existen bases de datos con capas de información, algunas a 1 km de resolución y que podrían ser de utilidad para cierto tipo de cultivos, un ejemplo es soilgrids (http://soilgrids.org/) la cual tiene ciertas propiedades del suelo estimadas a partir de algoritmos de machine

learning. Otra posibilidad es recolectar análisis químicos de suelo ya realizados o aplicar métodos más económicos como RASTA¹ (https://cgspace.cgiar.org/handle/10568/69682) el cual es un proceso sencillo que permite evaluar propiedades físicas del suelo.

Datos sobre las prácticas de manejo, en la mayoría de sistemas productivos en Latinoamérica y a la fecha (2016), normalmente no se tienen bases de datos previas, sólo en contados casos o a nivel de fincas aisladas. Es importante gestionar habilidades en toma de registros por parte de los productores o asegurar una red de técnicos o profesionales que actúen como intermediarios de información.

Definir estrategia para colectar información primaria

Se debe diseñar un sistema de colecta de datos, predio a predio o mejor lote a lote. Existen diferentes alternativas, cada una con puntos a favor y en contra: encuestas en papel, encuestas usando instrumentos de cómputo, tabletas o móviles, en estructuras básicas utilizando Excel o elaboradas en aplicaciones virtuales, off-line u on-line; una de las opciones que se pone a disposición, es la aplicación y/o plataforma AEPS: (https://play.google.com/store/apps/details?id=com.aepsmovil. aepsmovil&hl=es), que fueron diseñadas para captura y reporte de datos comerciales de maíz y fríjol. (Anexo 2).

Cada finca se debe subdividir por lotes y estos serán las unidades de observación; los lotes se reconocen como un área de cultivo definida por el agricultor que es homogénea en términos ambientales, lo cual sugiere que dentro de esta área, tanto el clima como el suelo varían muy poco, además se caracteriza porque el manejo que le da el agricultor es uniforme. Una finca podría tener un solo lote sí las condiciones de todo su cultivo son muy parecidas entre sí y se le da el mismo manejo.

Por cada lote se deben recopilar sus respectivas variables, de manera que se esperaría que a cada uno se le asocie un rendimiento, una caracterización del suelo, las condiciones climáticas y las prácticas de manejo. De igual forma la información que genere AEPS, deberá ser a nivel de lote.

Es recomendable que cada lote se asocie un nombre, puede ser combinado con números, que servirá de identificador único y corresponderá a los registros que se hagan al almacenar en la base de datos, facilitando la búsqueda y el seguimiento en el tiempo.

Existe la posibilidad que en el futuro por alguna razón se pierdan algunos datos ya sea porque no se anotó el registro por olvido o se dañó el dispositivo de medida, o los apuntes no son legibles por quien va a digitalizar; para esto es ideal crear una estrategia de indicación de datos faltantes. De las experiencias más problemáticas es cuando se opta por simplemente dejar vacío sin ninguna aclaración, puesto que se confunde con una ausencia, por ejemplo para el caso de fertilización un vacío puede representar como sí no hubiera habido aplicación o como sí se tratara de un dato faltante y esto repercute directamente en los resultados, por tanto se recomienda nunca dejar espacios vacíos en la base de datos y utilizar una nomenclatura como la siguiente: valores ceros donde no hubo aplicación, **NA** cuando no hay sentido que el campo contenga un valor y **ND** (No hay dato) cuando el dato está perdido.

Construir un diccionario de datos:

Se debe construir un conjunto de metadatos donde se describen las características de cada variable, en este caso se parte de la lista de variables identificadas en los pasos anteriores; a cada una se le debe asignar un nombre corto único el cual será la fila de encabezados en la base de datos, no se debe utilizar espacios ni comenzar por números o símbolos, ni contener tildes y es mucho mejor si no supera los 15 caracteres sin embargo hay casos en que es inevitable, se pueden separar palabras a partir de mayúsculas, algunos ejemplos son: CantFertQuimica (Cantidad de fertilización química), ProfEfecSuelo Profundidad efectiva del suelo o FechaSiembra (Fecha de siembra).

Es de gran relevancia construir un documento aparte donde se especifica las características mínimas para poder comprender la base de datos, este sería el diccionario de datos. Puede ser una tabla sencilla en la que se incluyen los siguientes campos: *clasificación* en la cual se define si se está hablando de una variable de suelo, manejo agronómico, clima,..., etc; *nombre corto* que sería el campo con las características nombradas

al inicio de esta sección, *descripción* o nombre completo de la variable, *unidad* de medida de la variable que se especificaría como: categórica en caso de que sea una variable con clases (Ej. CultivoAnterior, variedad, NompreProductor...), para las fechas se indicaría el formato, (dd/mm/yyyy, yyyy-mm-dd,...) y para variables numéricas, la unidad métrica (kilogramos, metros, hectárea,...). También se debe adicionar un campo llamado *rango de variación* el cual se define para variables numéricas: como el mínimo y máximo valor que puede tomar y para variables cualitativas todas las categorías posibles; en el caso de fechas, la fecha más antigua y la más reciente. Hay algunas variables que no tiene sentido calcular un rango de variación como cedula o nombres, en tal caso solo se deja indicada con un punto. La tabla muestra un ejemplo con pequeños fragmentos de un diccionario de datos.

Tabla 1. Ejemplo de diccionario de datos

Clasificación	Nombre Corto	Descripción	Unidad	Rango de variación	
General	NombreProductor	Nombre del producto	Categorica	•	
				Manual,	
Manejo	SistemaSiembra	Sistema de siembra	Categorica	Mecanizado,	
				Otro	
Manejo	FechaSiembra	Fecha de siembra	(dd/mm/yyyy)	(04/01/2010 -	
iviariejo	rechasiembra	recha de siembra (dd/mm/yyyy)		31/12/2016)	
Clima	Tmax	Temperatura	Grados Celsius	15 - 40	
Cilitia	IIIIdX	máxima	Grauos Ceisius	15 - 40	
Suelos	рН	pH del suelo	•	0-14	
Dondinsianta	Dondinsionto	Rendimiento	//:// anama a // a at 6 ma a /	(500 15000)	
Rendimiento	Rendimiento	obtenido	(kilogramos/hectárea)	(500 - 15000)	

Almacenamiento de datos

El almacenamiento de la información es de vital importancia ya que es lo que garantiza que se tendrá los datos disponibles y organizados en el momento en el que se requieren, se pueden utilizar diferentes gestores de bases de datos disponibles, como: MySQL, Microsoft Access u Oracle, los cuales facilitan la consulta, almacenamiento y administración de la información, para esto se debe construir una base de datos relacional, las plataformas de registros como AEPS normalmente ya la tienen integrada y se pueden descargar las consultas desde allí. No obstante alguien que aún no quiere comenzar a implementar un gestor de bases de datos como este, por los costos que acarrea o por la complejidad; puede organizar una base de datos desde Excel muy bien estructurada, teniendo en cuenta los identificadores o llaves que conectan las tablas y una vez tenga un buen volumen de información depurada con el tiempo, puede migrar su información a gestor de base de datos. Por último, algunos muy buenos consejos de administración y organización de bases de datos se pueden consultar aquí (http://www.decodigo.com/2012/02/buenas-practicas-para-el-uso-de-bases.html).

Anexo 1. Variables colectadas (Ejemplo arroz)

Características generales del lote

Departamento		
Municipio		
Finca		
Agricultor		
Nombre del lote		
Lat	decimales	Tomar un punto relativamente central en el lote
Long	decimales	Tomai un punto relativamente central en enote
Alt	msnm	
Area	На	
Propiedad	Lista	Propietario, arrienda, préstamo sin contrato

Variables de manejo en ciclo productivo

Obj	etivo			Practica	
		Dato de la practica	Tipo	Opciones pe	ensadas
		Fecha de trabajo	Fecha	_	
tivo	Preparacion de la parcela	tipo de preparación	Lista	Arada, Rastreada, Rastrillada, Caballoneo tradicional, caballoneao taipa, Nivelación pala, nivelación land plane, Fangueado, Quema, Drenajes	
Establecimiento del cultivo		Profundidad de trabajo (para labores que lo justifican)	Numero	(cm)	
ento		Fecha de siembra	Fecha	•	
mi				Secano,	
stableci		Sistema de cultivo Método de siembra	Lista Lista	riego, otro Trasplante, directa, voleo	
ŭ		Cantidad de semillas /			
	Siembra	ha Tipo de material	Número Lista	Kg/ha Variedad, Hibrido, OGM, semilla campesina	
		Material genético (nombre) Semillas tratadas ?	Lista Lista	Lista de los materiales usados en Colombia (los mas sembrados y otros) SI/NO	

		1		- Funcicidos		I		I
		Con que producto Objetivo de rendimiento	Lista	Fungicidas, insecticidas, otro (kg/ha) cuánto espero del cultivo ?				
	Datos generales	Cultivo anterior	Numero Lista	Lista de cultivos de Colombia ninguno, quema,	Soya, arroz, algodó n, maíz, sorgo, pastos, otros			
		Manejo de rastrojos	Lista	integración al suelo, picados (desbrozadora o combinada)				
		Fecha de aplicación	Fecha	0 ′ :		0 / :		
		Tipo de abono	Lista	Químico		Orgánico	Enmi	endas
c	Fertilización	Nombre del producto	Lista		Lista	Lista grandes categorías (Estiércol, pollinaza, gallinaza, compost de pollinaza, compost de gallinaza, compost vegetal, vinazas))	Lista	Lista principale s fuentes (Escorias, cal dolomita, yeso)
Nutrición		Cantidad (de producto bruto) kg/ha	Numero	Kg/ha N, P, K, Ca, Mg, S, Zn, Si	Numer o	Kg/ha	Numero	Kg/ha
		Composición (%) Cantidad de	Numero	%, %, %, %, %, %, %, %				
		elementos (kg) Forma de aplicación	Numero Lista	Foliar, edáfico				
						1		
	0	Fecha Cantidad aportada	Fecha					
	Riego	por hectárea (lamina)	Numero	(mm/ha) Aspersión,				
		Tipo de riego	Lista	gravedad.		J		
Protección	Control de Plagas,	Fecha del control	Fecha	Plagas,				
Prof	Cor	Tipo de objetivo	Lista	enfermedade s, arvenses.				

	Objetivo del control	Lista	Lista de plagas, enfermedade s (más comunes) y arvenses del arroz, otro. (3listas)		
		Químico	Familia de molécula activa	Lista Numer	Lista de moléculas usadas en Colombia Varias opciones de unidades
		Biológico	Tipo	Lista	Hongos antagonistas, Insectos, Bacterias, virus, arañas, feromonas, entomopatógeno s
			Dosis	Numer o	Varias opciones de unidades (L/ha, g/ha, individuos/ha)

	Fecha del monitoreo Fisiología	Fecha		PyE y arve	nses
Monitoreo	Fecha de emergencia Población a los 20 días Fecha de floración	Fecha Numero Fecha	Principales plagas, enfermedade s y arvenses (las dos principales) Incidencia	Lista	Listas de las 20 principales Plagas, Enfermedades, y Arvenses %

na L	Fecha de cosecha	Fecha	
Cosecha	Método	Lista	combinada, manual
ión	Cantidad (de todo el lote)	Numero	Toneladas
Producción	Rendimiento	Numero	kg/ha
po	% humedad	Numero	%
P	Almacenamiento	SI/NO	

Variables de suelo

Variables
Textura
%A
%Ar
%L
рH
%MO
Profundidad efectiva
Pendiente
Drenaje interno
Drenaje externo
Salinidad
Nutrientes
N
Р
K
S
Ca
Mg
Na
Cu
Fe
Zn
Mn
В

Variables de clima

Variables	Unidad	Resolución
Temperatura máxima	Grados Celsius	Ideal diario
Temperatura mínima	Grados Celsius	Ideal diario
Precipitación	mm	Ideal diario
Humedad relativa	%	Ideal diario
Radiación solar	Watts.m2, MJ.m-2, or cal.cm_2	Ideal diario
Brillo Solar	Horas luz	Diario

ENCUESTAS AGRICULTURA ESPECIFICA POR SITIO – AEPS FENALCE – CIAT



PRODUCTOR	
Nombre del productor:	
Tipo de documento:	○Cedula ○Pasaporte ○NIT ○Otro:
Numero de documento:	
Dirección:	
Departamento:	
Municipio:	
Teléfono fijo:	
Celular:	
Correo electrónico:	
FINCA	
Nombre de Finca:	
Latitud de la finca (°/decimale	-
Longitud de la Finca (°/decima	les):
Altitud de la Finca (metros):	
Departamento:	
Municipio:	
Vereda:	
Indicación (Como llegar):	
LOTE	
Tenencia del lote:	○Propio ○Arrendado ○Uso sin contrato
Nombre de Lote:	
Latitud del lote (°/decimales):	
Longitud del Lote (°/decimales):
Altitud del Lote (metros):	
Area del Lote (hectáreas):	
CULTIVO	
Cultivo anterior:	
Fecha de manejo del rastrojo:	
Manejo del rastrojo:	Quema OIntegrado al suelo OPicado ORetirado del lote
Tipo de preparación	○ Subsolador ○ Arado ○ Cincel ○ Rastra ○ Otro:
Profundidad del trabajo (cm):	
Numero de pases:	
Fecha de la labor:	
Fecha de siembra:	
Tipo de siembra:	OMecanizada OManual
Mejor rendimiento obtenido	
del lote (Kg/ha):	

Se hace drenaje a la parcela:		OSi ONe	•			
Kg. de semilla utilizada por ha:						
Semillas tratadas:	OSi ONe	Pro	ducto:			
Distancia entre surcos (metros):						
Distancia entre sitios (metros):						
Numero de semillas por sitio:						
Tipo de material:		○ Variedad	OHil	brido ()Transgénio	со
Color del endospermo:						
Material genético (nombre):						
Fecha de emergencia:						
Población a los 20 días (# plantas	s):					
Fecha de floración:						
Porcentaje de resiembra:						·
Tipo de riego:		Aspersión	O	iravedad	○Por go	teo
Fecha(s) del riego:						
Cantidad aportada por hectárea	(m³):					
FERTILIZACION						
Tipo de Fertilizante:	()Orga	ánico ()Qu	ímico			
Fecha de aplicación:						
Tipo de aplicación:	()Edát	fico ()Folia	г			
Producto:						
Cantidad:			Uni	dades:	○Kg/ha	OL/ha
Tipo de Fertilizante:	Orga	ánico ()Qu	ímico			
Fecha de aplicación:						
Tipo de aplicación:	()Edát	fico ()Folia	г			
Producto:						
Cantidad:			Unio	dades:	○Kg/ha	OL/ha
Tipo de Fertilizante:	()Orga	ánico ()Qu	ímico			
Fecha de aplicación:						
Tipo de aplicación:	⊜Edáfico ⊜Foliar					
Producto:						
Cantidad:	Unidades: ()Kg/ha ()L/ha					
Enmiendas:	⊘ Si	()No				
Fecha de aplicación:						
Producto:						
Cantidad:			Uni	dades:	○Kg/ha	OL/ha
L '	•					٠.

Fecha del monitoreo:		
Monitoreo:	○Plag	as (Enfermedades (Malezas
Objetivo del monitoreo:		
Porcentaje de incidencia:		
Fecha del monitoreo:		
Monitoreo:	○Plag	as (Enfermedades ()Malezas
Objetivo del monitoreo:		
Porcentaje de incidencia:		
Fecha del control:		
Control de:		OPlagas OMalezas OEnfermedades
Objetivo del control (cogollero, por ej):		
Tipo de control:		OBiológico OQuímico OMecánico
Ingrediente activo:		
Dosis:		Unidad:
Fecha del control:		
Control de:		Plagas Malezas Enfermedades
Objetivo del control (cogollero, por ej):		
Tipo de control:		OBiológico OQuímico OMecánico
Ingrediente activo:		
Dosis:		Unidad:
		T
Fecha del control:		
Control de:		Plagas Malezas Enfermedades
Objetivo del control (cogollero, por ej):		0
Tipo de control:		OBiológico OQuímico OMecánico
Ingrediente activo:		I
Dosis:		Unidad:
Fecha del control:		
Control de:		○Plagas ○Malezas ○Enfermedades
		Plagas Malezas Enfermedades
Objetivo del control (cogollero, por ej):		Biológico Químico Mecánico
Tipo de control:		Biológico Químico Mecánico
Ingrediente activo:		Unidad:
Dosis:		Onidad:
Fecha del control:		
Control de:		○Plagas ○Malezas ○Enfermedades
Objetivo del control (cogollero, por ej):		Charles Chinesticanes
Tipo de control:		○Biológico ○Químico ○Mecánico
Ingrediente activo:		Conseque (Administra (Association
Dosis:		Unidad:
e costs.		onidad.

COSECHA

Fecha de cosecha:		
Método de cosecha:	OMecanizada OManual	
Producto cosechado:	○Grano seco ○Mazorca fresca ○Ensilaje	
	-	
Rendimiento en Kg/ha (solo para grano seco):		
% de Humedad de la cosecha (solo para grano seco):		
Numero de buitos/ha (solo para mazorca fresca):		
Peso promedio de bulto en Kg/bulto (solo para mazorca fresca):		
Numero de bolsas/ha (solo para ensilaje):		
Peso promedio por bolsa en Kg/bolsa (solo para ensilaje):		
Cantidad total en Kg (Kg	/ha multiplicado	
por número de hectárea	s):	
Almacenamiento en fino	a: OSi ONo	
Observaciones que afec	taron la	
producción:		