

# faoswsSeed: A package for the imputation of the seed domain of the Statistical Working System

Joshua M. Browning, Michael. C. J. Kao

Food and Agriculture Organization  
of the United Nations

---

## Abstract

This vignette provides detailed description of the usage of functions in the **faoswsSeed** package.

There are two sections to this paper. The first is introductory, and provides a brief overview of the algorithm followed by this package. The second section shows a sample execution of the module, and describes what each function is doing as execution proceeds.

*Keywords:* Seed, Agriculture.

---

## 1. Introduction

This algorithm follows this general process:

1. Pull agricultural data from the database.
2. Estimate the area sown for each year.
3. Estimate the seed rate.
4. Estimate the total seed used by multiplying the area sown by the seed rate.
5. Push the updated data.table back to the database.

## 2. Example

Before we begin, we will need to load the required library

```
> ## Load libraries
> library(faosws)
> library(faoswsSeed)
> library(faoswsImputation)
> library(faoswsUtil)
> library(data.table)
> library(ggplot2)
> library(faoswsProcessing)
>
```

## 2.1. Pull Data

Now, we need to get a `data.table` object from the working system. To do this, we'll need a token for the R session. To get this token, you'll need to create an `.xml` file that references the appropriate dataset: see the `.xml` file in the `sws_seed` repository. Alternatively, you can create an `.xml` file with the Code tag set to agriculture and four `SelectableDimension`'s: `geographicAreaM49`, `measuredItemCPC`, `measuredElement`, `timePointYears`. Once you get this token, you can run `GetTestEnvironment` (see below, `baseUrl` is the same) to load some variables (usually starting with `swsContext.*`) into your workspace.

Now, we may change the dataset of interest by updating the keys in the `swsContext.datasets[[1]]` object. For example, let's look at Austria's data (`geographicAreaM49` code of "40"). Note: multiple keys are allowed.

Lastly, the `getAreaData` function will pull in the seed data needed. The code below shows an example of what these calls look like, although it is not executable as the token is no longer valid.

```
> ## Pull in necessary swsContext parameters, see faosws documentation
> GetTestEnvironment(
+   baseUrl = "https://hqlqasws1.hq.un.fao.org:8181/sws",
+   token = "ec5a4b0e-0ffa-432e-9db0-ba08072c924b"
+ )
> swsContext.datasets[[1]]@dimensions$geographicAreaM49@keys = c("40")
> data = getAreaData(dataContext = swsContext.datasets[[1]],
+   areaSownElementCode = "5212",
+   areaHarvestedElementCode = "5312",
+   seedElementCode = "5525")
```

We can't pull this data directly, but the `faoswsSeed` package has a default dataset that would result from this kind of a call:

```
> head(seedData)
```

	geographicAreaM49	measuredItemCPC	timePointYears	Value_measuredElement_5212
1:	100	0111	1994	NA
2:	100	0111	1995	NA
3:	100	0111	1996	NA
4:	100	0111	1997	NA
5:	100	0111	1998	NA
6:	100	0111	1999	NA

  

	flagObservationStatus_measuredElement_5212	flagMethod_measuredElement_5212
1:		M
2:		M
3:		M
4:		M
5:		M
6:		M

  

	Value_measuredElement_5312	flagObservationStatus_measuredElement_5312
1:	1319760	
2:	1181120	
3:	957670	
4:	1211720	

```

5:          1141682
6:          966282
  flagMethod_measuredElement_5312 Value_measuredElement_5525
1:          -          NA
2:          -          NA
3:          -          NA
4:          -          NA
5:          -          NA
6:          -          NA
  flagObservationStatus_measuredElement_5525 flagMethod_measuredElement_5525
1:          M          u
2:          M          u
3:          M          u
4:          M          u
5:          M          u
6:          M          u

```

The data is somewhat cryptic, as there are a lot of codes used, but it's there.

## 2.2. Estimate Area Sown

Next, we need to impute the missing values. Imputation follows one of several methods:

1. If no area sown values exist, the area sown is imputed as the area harvested.
2. If some area sown values exist, then imputation is performed based on input parameters. If the `imputationParameters` argument is `NULL`, then an average ratio is computed across all non-missing values within each `byKey` group (by default, `byKey` is `NULL` and so the ratio is computed with all the data):  $R = (\text{area sown}) / (\text{area harvested})$ . Missing values for area sown are then imputed by taking the area harvested and multiplying by this ratio.
3. If some area sown values exist and `imputationParameters` is not `NULL`, then imputation is performed via ensemble imputation from the `faoswsImputation` package. All of the elegant models in that framework are not likely to be useful here, as most countries have no data. However, two models such as a local and global mean may be helpful.

In this example, we have the first scenario (remember, 5212 is the area sown code and 5312 is the area harvested).

```

      time areaSown areaHarvested
1: 1994    131916    101212
2: 1995    131334    100108
3: 1996    130934     99660
4: 1997    130874     98901
5: 1998    129658     99099
6: 1999    127066     99000
7: 2000    106000     88672
8: 2001     93000     82186
9: 2002     93000     82846
10: 2003     93100     93032
11: 2004     93000     93217
12: 2005    100000     86028

```

13: 2006	80000	75634
14: 2007	80000	75260
15: 2008	83000	75776
16: 2009	83000	75933
17: 2010	NA	73922
18: 2011	83000	75511
19: 2012	NA	NA
20: 2013	NA	NA

	geographicAreaM49	measuredItemCPC	timePointYears
1:	100	0111	1994
2:	100	0111	1995
3:	100	0111	1996
4:	100	0111	1997
5:	100	0111	1998
---			
8956:	348	26190.01	2009
8957:	348	26190.01	2010
8958:	348	26190.01	2011
8959:	348	26190.01	2012
8960:	348	26190.01	2013

  

	Value_measuredElement_5212	flagObservationStatus_measuredElement_5212
1:	1330631.1	I
2:	1190849.1	I
3:	965558.5	I
4:	1221701.2	I
5:	1151086.2	I
---		
8956:	NA	M
8957:	NA	M
8958:	NA	M
8959:	NA	M
8960:	NA	M

  

	flagMethod_measuredElement_5212	Value_measuredElement_5312
1:	e	1319760
2:	e	1181120
3:	e	957670
4:	e	1211720
5:	e	1141682
---		
8956:	e	NA
8957:	e	NA
8958:	e	NA
8959:	e	NA
8960:	e	NA

  

	flagObservationStatus_measuredElement_5312
1:	
2:	
3:	
4:	
5:	
---	

```

8956:                                     M
8957:                                     M
8958:                                     M
8959:                                     M
8960:                                     M
flagMethod_measuredElement_5312 Value_measuredElement_5525
  1:                                     -      NA
  2:                                     -      NA
  3:                                     -      NA
  4:                                     -      NA
  5:                                     -      NA
---
8956:                                     u      NA
8957:                                     u      NA
8958:                                     u      NA
8959:                                     u      NA
8960:                                     u      NA
flagObservationStatus_measuredElement_5525
  1:                                     M
  2:                                     M
  3:                                     M
  4:                                     M
  5:                                     M
---
8956:                                     M
8957:                                     M
8958:                                     M
8959:                                     M
8960:                                     M
flagMethod_measuredElement_5525 Value_areaSownRatio
  1:                                     u      1.008237
  2:                                     u      1.008237
  3:                                     u      1.008237
  4:                                     u      1.008237
  5:                                     u      1.008237
---
8956:                                     u      1.008237
8957:                                     u      1.008237
8958:                                     u      1.008237
8959:                                     u      1.008237
8960:                                     u      1.008237
flagObservationStatus_areaSownRatio flagMethod_areaSownRatio
  1:                                     I      e
  2:                                     I      e
  3:                                     I      e
  4:                                     I      e
  5:                                     I      e
---
8956:                                     I      e
8957:                                     I      e
8958:                                     I      e

```

```

8959:                                I                                e
8960:                                I                                e

```

	time	areaSown	areaHarvested	Value_areaSownRatio
1:	1994	131916.00	101212	1.008237
2:	1995	131334.00	100108	1.008237
3:	1996	130934.00	99660	1.008237
4:	1997	130874.00	98901	1.008237
5:	1998	129658.00	99099	1.008237
6:	1999	127066.00	99000	1.008237
7:	2000	106000.00	88672	1.008237
8:	2001	93000.00	82186	1.008237
9:	2002	93000.00	82846	1.008237
10:	2003	93100.00	93032	1.008237
11:	2004	93000.00	93217	1.008237
12:	2005	100000.00	86028	1.008237
13:	2006	80000.00	75634	1.008237
14:	2007	80000.00	75260	1.008237
15:	2008	83000.00	75776	1.008237
16:	2009	83000.00	75933	1.008237
17:	2010	74530.91	73922	1.008237
18:	2011	83000.00	75511	1.008237
19:	2012	NA	NA	1.008237
20:	2013	NA	NA	1.008237

The areaSownRatio is estimated globally by default. We could also estimate the ratio within each country and commodity individually:

```

> temp = copy(seedData)
> imputeAreaSown(data = temp, codeAreaSown = "5212",
+               byKey = c("geographicAreaM49", "measuredItemCPC"))

```

	geographicAreaM49	measuredItemCPC	timePointYears
1:	100	0111	1994
2:	100	0111	1995
3:	100	0111	1996
4:	100	0111	1997
5:	100	0111	1998

---

8956:	348	26190.01	2009
8957:	348	26190.01	2010
8958:	348	26190.01	2011
8959:	348	26190.01	2012
8960:	348	26190.01	2013

	Value_measuredElement_5212	flagObservationStatus_measuredElement_5212
--	----------------------------	--

1:	1319760
2:	1181120
3:	957670
4:	1211720
5:	1141682

---

8956:	NA	M
8957:	NA	M
8958:	NA	M
8959:	NA	M
8960:	NA	M

flagMethod\_measuredElement\_5212 Value\_measuredElement\_5312

1:	e	1319760
2:	e	1181120
3:	e	957670
4:	e	1211720
5:	e	1141682

----

8956:	e	NA
8957:	e	NA
8958:	e	NA
8959:	e	NA
8960:	e	NA

flagObservationStatus\_measuredElement\_5312

1:		
2:		
3:		
4:		
5:		

----

8956:	M
8957:	M
8958:	M
8959:	M
8960:	M

flagMethod\_measuredElement\_5312 Value\_measuredElement\_5525

1:	-	NA
2:	-	NA
3:	-	NA
4:	-	NA
5:	-	NA

----

8956:	u	NA
8957:	u	NA
8958:	u	NA
8959:	u	NA
8960:	u	NA

flagObservationStatus\_measuredElement\_5525

1:	M
2:	M
3:	M
4:	M
5:	M

----

8956:	M
8957:	M
8958:	M

```

8959:                                     M
8960:                                     M
      flagMethod_measuredElement_5525 Value_areaSownRatio
1:                                     u      NaN
2:                                     u      NaN
3:                                     u      NaN
4:                                     u      NaN
5:                                     u      NaN
---
8956:                                     u      NaN
8957:                                     u      NaN
8958:                                     u      NaN
8959:                                     u      NaN
8960:                                     u      NaN
      flagObservationStatus_areaSownRatio flagMethod_areaSownRatio
1:                                     I      e
2:                                     I      e
3:                                     I      e
4:                                     I      e
5:                                     I      e
---
8956:                                     I      e
8957:                                     I      e
8958:                                     I      e
8959:                                     I      e
8960:                                     I      e

> temp[geographicAreaM49 == 348 & measuredItemCPC == "01330",
+       .(time = timePointYears, areaSown = Value_measuredElement_5212,
+       areaHarvested = Value_measuredElement_5312, Value_areaSownRatio)]

      time areaSown areaHarvested Value_areaSownRatio
1: 1994 131916.00      101212      1.168406
2: 1995 131334.00      100108      1.168406
3: 1996 130934.00      99660      1.168406
4: 1997 130874.00      98901      1.168406
5: 1998 129658.00      99099      1.168406
6: 1999 127066.00      99000      1.168406
7: 2000 106000.00      88672      1.168406
8: 2001 93000.00      82186      1.168406
9: 2002 93000.00      82846      1.168406
10: 2003 93100.00      93032      1.168406
11: 2004 93000.00      93217      1.168406
12: 2005 100000.00      86028      1.168406
13: 2006 80000.00      75634      1.168406
14: 2007 80000.00      75260      1.168406
15: 2008 83000.00      75776      1.168406
16: 2009 83000.00      75933      1.168406
17: 2010 86370.91      73922      1.168406
18: 2011 83000.00      75511      1.168406
19: 2012      NA      NA      1.168406
20: 2013      NA      NA      1.168406

```



Or, we can estimate the ratio via an ensemble method (from faoswsImputation):

```
> imputationParams = defaultImputationParameters(variable = "seed")
> ## Coerce type to character instead of default factor type
> imputationParams$flagTable$flagObservationStatus =
+   as.character(imputationParams$flagTable$flagObservationStatus)
> imputationParams$ensembleModels = list(
+   defaultMean = ensembleModel(model = defaultMean, extrapolationRange = Inf,
+                               level = "local"),
+   globalMean = ensembleModel(model = defaultGlobalMean,
+                               extrapolationRange = Inf, level = "global"),
+   defaultLm = ensembleModel(model = defaultLm, extrapolationRange = Inf,
+                              level = "local"))
> temp = seedData[measuredItemCPC == "01330", ]
> imputeAreaSown(data = temp, codeAreaSown = "5212",
+   imputationParameters = imputationParams)
```

	geographicAreaM49	measuredItemCPC	timePointYears
1:	100	01330	1994
2:	100	01330	1995
3:	100	01330	1996
4:	100	01330	1997
5:	100	01330	1998
6:	100	01330	1999
7:	100	01330	2000
8:	100	01330	2001
9:	100	01330	2002
10:	100	01330	2003
11:	100	01330	2004
12:	100	01330	2005
13:	100	01330	2006
14:	100	01330	2007
15:	100	01330	2008
16:	100	01330	2009
17:	100	01330	2010
18:	100	01330	2011
19:	100	01330	2012
20:	100	01330	2013
21:	348	01330	1994
22:	348	01330	1995
23:	348	01330	1996
24:	348	01330	1997
25:	348	01330	1998
26:	348	01330	1999
27:	348	01330	2000
28:	348	01330	2001
29:	348	01330	2002
30:	348	01330	2003
31:	348	01330	2004
32:	348	01330	2005
33:	348	01330	2006
34:	348	01330	2007

35:	348	01330	2008
36:	348	01330	2009
37:	348	01330	2010
38:	348	01330	2011
39:	348	01330	2012
40:	348	01330	2013
41:	400	01330	1994
42:	400	01330	1995
43:	400	01330	1996
44:	400	01330	1997
45:	400	01330	1998
46:	400	01330	1999
47:	400	01330	2000
48:	400	01330	2001
49:	400	01330	2002
50:	400	01330	2003
51:	400	01330	2004
52:	400	01330	2005
53:	400	01330	2006
54:	400	01330	2007
55:	400	01330	2008
56:	400	01330	2009
57:	400	01330	2010
58:	400	01330	2011
59:	400	01330	2012
60:	400	01330	2013
61:	638	01330	1994
62:	638	01330	1995
63:	638	01330	1996
64:	638	01330	1997
65:	638	01330	1998
66:	638	01330	1999
67:	638	01330	2000
68:	638	01330	2001
69:	638	01330	2002
70:	638	01330	2003
71:	638	01330	2004
72:	638	01330	2005
73:	638	01330	2006
74:	638	01330	2007
75:	638	01330	2008
76:	638	01330	2009
77:	638	01330	2010
78:	638	01330	2011
79:	638	01330	2012
80:	638	01330	2013
81:	804	01330	1994
82:	804	01330	1995
83:	804	01330	1996
84:	804	01330	1997
85:	804	01330	1998

86:	804	01330	1999
87:	804	01330	2000
88:	804	01330	2001
89:	804	01330	2002
90:	804	01330	2003
91:	804	01330	2004
92:	804	01330	2005
93:	804	01330	2006
94:	804	01330	2007
95:	804	01330	2008
96:	804	01330	2009
97:	804	01330	2010
98:	804	01330	2011
99:	804	01330	2012
100:	804	01330	2013

geographicAreaM49 measuredItemCPC timePointYears

Value\_measuredElement\_5212 flagObservationStatus\_measuredElement\_5212

1:	133408.46265	I
2:	131806.80738	I
3:	126013.76158	I
4:	125894.81513	I
5:	132381.51898	I
6:	130223.99511	I
7:	130967.11605	I
8:	151588.42768	I
9:	134998.34105	I
10:	154358.34915	I
11:	152519.97865	I
12:	149380.26324	I
13:	151753.30395	I
14:	141724.11550	I
15:	130506.64016	I
16:	119457.57416	I
17:	97365.33059	I
18:	92410.79844	I
19:	NA	M
20:	NA	M
21:	131916.00000	
22:	131334.00000	
23:	130934.00000	
24:	130874.00000	
25:	129658.00000	
26:	127066.00000	
27:	106000.00000	
28:	93000.00000	
29:	93000.00000	
30:	93100.00000	
31:	93000.00000	
32:	100000.00000	E
33:	80000.00000	E
34:	80000.00000	E

35:	83000.00000	
36:	83000.00000	
37:	81618.23183	I
38:	83000.00000	E
39:	NA	M
40:	NA	M
41:	2356.00000	
42:	2385.00000	
43:	2424.00000	
44:	3684.00000	
45:	3705.00000	
46:	3709.00000	
47:	3739.00000	
48:	3771.00000	
49:	3904.00000	
50:	3630.00000	
51:	3635.00000	
52:	3635.00000	
53:	3646.00000	
54:	3089.00000	
55:	3110.00000	
56:	3138.00000	
57:	3199.00000	
58:	3941.00000	
59:	NA	M
60:	NA	M
61:	23.55375	I
62:	23.55375	I
63:	23.55375	I
64:	27.08682	I
65:	29.44219	I
66:	28.26450	I
67:	29.44219	I
68:	30.61988	I
69:	31.79757	I
70:	32.97526	I
71:	23.55375	I
72:	23.55375	I
73:	34.15294	I
74:	23.55375	I
75:	NA	M
76:	24.73144	I
77:	35.33063	I
78:	36.50832	I
79:	NA	M
80:	NA	M
81:	160100.00000	
82:	155000.00000	
83:	148200.00000	
84:	138600.00000	
85:	126000.00000	

86:	116000.00000	
87:	110000.00000	
88:	105000.00000	
89:	103000.00000	
90:	99000.00000	
91:	97000.00000	
92:	96000.00000	
93:	93000.00000	
94:	93000.00000	
95:	93000.00000	T
96:	93000.00000	T
97:	88000.00000	T
98:	90200.00000	T
99:	NA	M
100:	NA	M

  

	Value_measuredElement_5212	flagObservationStatus_measuredElement_5212	flagMethod_measuredElement_5212	Value_measuredElement_5312
1:	e			113280
2:	e			111920
3:	e			107001
4:	e			106900
5:	e			112408
6:	e			110576
7:	e			111207
8:	e			128717
9:	e			114630
10:	e			131069
11:	e			129508
12:	e			126842
13:	e			128857
14:	e			120341
15:	e			110816
16:	e			101434
17:	e			82675
18:	e			78468
19:	e			NA
20:	e			NA
21:	-			101212
22:	-			100108
23:	-			99660
24:	-			98901
25:	-			99099
26:	-			99000
27:	-			88672
28:	-			82186
29:	-			82846
30:	-			93032
31:	-			93217
32:	f			86028
33:	f			75634
34:	f			75260

35:	-	75776
36:	-	75933
37:	e	73922
38:	f	75511
39:	e	NA
40:	e	NA
41:	e	2356
42:	e	2385
43:	e	2424
44:	e	3684
45:	e	3705
46:	e	3709
47:	e	3739
48:	e	3771
49:	e	3904
50:	e	3630
51:	e	3635
52:	e	3635
53:	e	3646
54:	e	3089
55:	e	3110
56:	e	3138
57:	e	3199
58:	e	3941
59:	e	NA
60:	e	NA
61:	e	20
62:	e	20
63:	e	20
64:	e	23
65:	e	25
66:	e	24
67:	e	25
68:	e	26
69:	e	27
70:	e	28
71:	e	20
72:	e	20
73:	e	29
74:	e	20
75:	e	NA
76:	e	21
77:	e	30
78:	e	31
79:	e	NA
80:	e	NA
81:	-	140000
82:	-	137700
83:	-	134323
84:	-	122900
85:	-	112300

86:	-	104800
87:	-	99400
88:	-	94700
89:	-	91200
90:	-	85500
91:	-	82800
92:	-	80600
93:	-	75800
94:	-	71200
95:	p	70900
96:	p	71000
97:	p	67600
98:	p	69100
99:	e	NA
100:	e	NA

flagMethod\_measuredElement\_5212 Value\_measuredElement\_5312  
flagObservationStatus\_measuredElement\_5312 flagMethod\_measuredElement\_5312

1:		-
2:		-
3:		-
4:		-
5:		-
6:		-
7:		-
8:		-
9:		-
10:		-
11:		-
12:		-
13:		-
14:		-
15:		-
16:		-
17:		-
18:		-
19:	M	u
20:	M	u
21:		-
22:		-
23:		-
24:		-
25:		-
26:		-
27:		-
28:		-
29:		-
30:		-
31:		-
32:		-
33:		-
34:		-

35: -  
36: -  
37: -  
38: -  
39: M u  
40: M u  
41: -  
42: -  
43: -  
44: -  
45: -  
46: -  
47: -  
48: -  
49: -  
50: -  
51: -  
52: -  
53: -  
54: -  
55: -  
56: -  
57: -  
58: -  
59: M u  
60: M u  
61: E f  
62: E f  
63: E f  
64: E e  
65: E e  
66: E e  
67: E e  
68: E e  
69: E e  
70: E e  
71: E f  
72: E f  
73: E e  
74: E f  
75: M u  
76: E e  
77: E f  
78: E e  
79: M u  
80: M u  
81: -  
82: -  
83: -  
84: -  
85: -



86:		-
87:		-
88:		-
89:		-
90:		-
91:		-
92:		-
93:		-
94:		-
95:		-
96:		-
97:		-
98:		-
99:	M	u
100:	M	u

flagObservationStatus\_measuredElement\_5312 flagMethod\_measuredElement\_5312  
 Value\_measuredElement\_5525 flagObservationStatus\_measuredElement\_5525

1:	NA	M
2:	NA	M
3:	NA	M
4:	NA	M
5:	NA	M
6:	NA	M
7:	NA	M
8:	NA	M
9:	NA	M
10:	NA	M
11:	NA	M
12:	NA	M
13:	NA	M
14:	NA	M
15:	NA	M
16:	NA	M
17:	NA	M
18:	NA	M
19:	NA	M
20:	NA	M
21:	NA	M
22:	NA	M
23:	NA	M
24:	NA	M
25:	NA	M
26:	NA	M
27:	NA	M
28:	NA	M
29:	NA	M
30:	NA	M
31:	NA	M
32:	NA	M
33:	NA	M
34:	NA	M

35:	NA	M
36:	NA	M
37:	NA	M
38:	NA	M
39:	NA	M
40:	NA	M
41:	NA	M
42:	NA	M
43:	NA	M
44:	NA	M
45:	NA	M
46:	NA	M
47:	NA	M
48:	NA	M
49:	NA	M
50:	NA	M
51:	NA	M
52:	NA	M
53:	NA	M
54:	NA	M
55:	NA	M
56:	NA	M
57:	NA	M
58:	NA	M
59:	NA	M
60:	NA	M
61:	NA	M
62:	NA	M
63:	NA	M
64:	NA	M
65:	NA	M
66:	NA	M
67:	NA	M
68:	NA	M
69:	NA	M
70:	NA	M
71:	NA	M
72:	NA	M
73:	NA	M
74:	NA	M
75:	NA	M
76:	NA	M
77:	NA	M
78:	NA	M
79:	NA	M
80:	NA	M
81:	NA	M
82:	NA	M
83:	NA	M
84:	NA	M
85:	NA	M

86:	NA	M
87:	NA	M
88:	NA	M
89:	NA	M
90:	NA	M
91:	NA	M
92:	NA	M
93:	NA	M
94:	NA	M
95:	NA	M
96:	NA	M
97:	NA	M
98:	NA	M
99:	NA	M
100:	NA	M

Value\_measuredElement\_5525 flagObservationStatus\_measuredElement\_5525  
 flagMethod\_measuredElement\_5525 Value\_areaSownRatio

1:	u	1.177688
2:	u	1.177688
3:	u	1.177688
4:	u	1.177688
5:	u	1.177688
6:	u	1.177688
7:	u	1.177688
8:	u	1.177688
9:	u	1.177688
10:	u	1.177688
11:	u	1.177688
12:	u	1.177688
13:	u	1.177688
14:	u	1.177688
15:	u	1.177688
16:	u	1.177688
17:	u	1.177688
18:	u	1.177688
19:	u	1.177688
20:	u	1.177688
21:	u	1.303363
22:	u	1.311923
23:	u	1.313807
24:	u	1.323283
25:	u	1.308368
26:	u	1.283495
27:	u	1.195417
28:	u	1.131580
29:	u	1.122565
30:	u	1.000731
31:	u	1.000000
32:	u	1.162412
33:	u	1.057725
34:	u	1.062982

35:	u	1.095334
36:	u	1.093069
37:	u	1.104113
38:	u	1.099178
39:	u	1.087283
40:	u	1.078869
41:	u	NA
42:	u	NA
43:	u	NA
44:	u	NA
45:	u	NA
46:	u	NA
47:	u	NA
48:	u	NA
49:	u	NA
50:	u	NA
51:	u	NA
52:	u	NA
53:	u	NA
54:	u	NA
55:	u	NA
56:	u	NA
57:	u	NA
58:	u	NA
59:	u	NA
60:	u	NA
61:	u	1.177688
62:	u	1.177688
63:	u	1.177688
64:	u	1.177688
65:	u	1.177688
66:	u	1.177688
67:	u	1.177688
68:	u	1.177688
69:	u	1.177688
70:	u	1.177688
71:	u	1.177688
72:	u	1.177688
73:	u	1.177688
74:	u	1.177688
75:	u	1.177688
76:	u	1.177688
77:	u	1.177688
78:	u	1.177688
79:	u	1.177688
80:	u	1.177688
81:	u	1.143571
82:	u	1.125635
83:	u	1.103311
84:	u	1.127746
85:	u	1.121995

86:	u	1.106870
87:	u	1.106640
88:	u	1.108765
89:	u	1.129386
90:	u	1.157895
91:	u	1.171498
92:	u	1.191067
93:	u	1.226913
94:	u	1.306180
95:	u	1.311707
96:	u	1.309859
97:	u	1.301775
98:	u	1.305355
99:	u	1.246365
100:	u	1.252916

flagMethod\_measuredElement\_5525 Value\_areaSownRatio

flagObservationStatus\_areaSownRatio flagMethod\_areaSownRatio

1:	M	e
2:	M	e
3:	M	e
4:	M	e
5:	M	e
6:	M	e
7:	M	e
8:	M	e
9:	M	e
10:	M	e
11:	M	e
12:	M	e
13:	M	e
14:	M	e
15:	M	e
16:	M	e
17:	M	e
18:	M	e
19:	M	e
20:	M	e
21:		e
22:		e
23:		e
24:		e
25:		e
26:		e
27:		e
28:		e
29:		e
30:		e
31:		e
32:	E	e
33:	E	e
34:	E	e

35:		e
36:		e
37:	M	e
38:	E	e
39:	M	e
40:	M	e
41:	M	e
42:	M	e
43:	M	e
44:	M	e
45:	M	e
46:	M	e
47:	M	e
48:	M	e
49:	M	e
50:	M	e
51:	M	e
52:	M	e
53:	M	e
54:	M	e
55:	M	e
56:	M	e
57:	M	e
58:	M	e
59:	M	e
60:	M	e
61:	M	e
62:	M	e
63:	M	e
64:	M	e
65:	M	e
66:	M	e
67:	M	e
68:	M	e
69:	M	e
70:	M	e
71:	M	e
72:	M	e
73:	M	e
74:	M	e
75:	M	e
76:	M	e
77:	M	e
78:	M	e
79:	M	e
80:	M	e
81:		e
82:		e
83:		e
84:		e
85:		e

```

86:                                     e
87:                                     e
88:                                     e
89:                                     e
90:                                     e
91:                                     e
92:                                     e
93:                                     e
94:                                     e
95:                                T     e
96:                                T     e
97:                                T     e
98:                                T     e
99:                                M     e
100:                               M     e
    flagObservationStatus_areaSownRatio flagMethod_areaSownRatio

> temp[geographicAreaM49 == 348 & measuredItemCPC == "01330",
+      .(time = timePointYears, areaSown = Value_measuredElement_5212,
+        areaHarvested = Value_measuredElement_5312, Value_areaSownRatio)]

   time  areaSown areaHarvested Value_areaSownRatio
1: 1994 131916.00      101212      1.303363
2: 1995 131334.00      100108      1.311923
3: 1996 130934.00       99660      1.313807
4: 1997 130874.00       98901      1.323283
5: 1998 129658.00       99099      1.308368
6: 1999 127066.00       99000      1.283495
7: 2000 106000.00       88672      1.195417
8: 2001  93000.00       82186      1.131580
9: 2002  93000.00       82846      1.122565
10: 2003  93100.00       93032      1.000731
11: 2004  93000.00       93217      1.000000
12: 2005 100000.00       86028      1.162412
13: 2006  80000.00       75634      1.057725
14: 2007  80000.00       75260      1.062982
15: 2008  83000.00       75776      1.095334
16: 2009  83000.00       75933      1.093069
17: 2010  81618.23       73922      1.104113
18: 2011  83000.00       75511      1.099178
19: 2012      NA        NA      1.087283
20: 2013      NA        NA      1.078869

```

In this example, notice that we filtered the seedData set to one specific measuredItemCPC code. This is a requirement for imputeAreaSown: it can only handle one CPC code at a time. A future goal of this package is to add an additional function that calls imputeAreaSown for each individual CPC code.

Additionally, note that the defaultLm model could be problematic: predictions in the later years for the left graph give ratios of less than 1. These don't make sense: you can't sow less area than your harvest. There is a check within the code that corrects any imputed values less than 1 (by setting it to 1); however, the analyst should ensure that the models they use do not impute values smaller than 1 as a good practice.

Note that this ensemble model seems to generate improved estimates of area sown rates in this scenario. However, in most scenarios, area sown is not available. Moreover, if it is, it is generally available for almost all years. The usual scenario would be more like the following example:

```
> data = seedData[measuredItemCPC == "0111", ]
> data[geographicAreaM49 == 100,
+      .(time = timePointYears, areaSown = Value_measuredElement_5212,
+      areaHarvested = Value_measuredElement_5312)]
```

	time	areaSown	areaHarvested
1:	1994	NA	1319760
2:	1995	NA	1181120
3:	1996	NA	957670
4:	1997	NA	1211720
5:	1998	NA	1141682
6:	1999	NA	966282
7:	2000	NA	978575
8:	2001	NA	1355500
9:	2002	NA	1368627
10:	2003	NA	841014
11:	2004	NA	1039680
12:	2005	NA	1101807
13:	2006	NA	970392
14:	2007	NA	1087996
15:	2008	NA	1111533
16:	2009	NA	1247718
17:	2010	NA	1137650
18:	2011	NA	1137642
19:	2012	NA	1090000
20:	2013	NA	NA

```
> imputeAreaSown(data = data, codeAreaSown = "5212",
+      imputationParameters = imputationParams)
```

	geographicAreaM49	measuredItemCPC	timePointYears
1:	100	0111	1994
2:	100	0111	1995
3:	100	0111	1996
4:	100	0111	1997
5:	100	0111	1998
6:	100	0111	1999
7:	100	0111	2000
8:	100	0111	2001
9:	100	0111	2002
10:	100	0111	2003
11:	100	0111	2004
12:	100	0111	2005
13:	100	0111	2006
14:	100	0111	2007
15:	100	0111	2008



16:	100	0111	2009
17:	100	0111	2010
18:	100	0111	2011
19:	100	0111	2012
20:	100	0111	2013
21:	894	0111	1994
22:	894	0111	1995
23:	894	0111	1996
24:	894	0111	1997
25:	894	0111	1998
26:	894	0111	1999
27:	894	0111	2000
28:	894	0111	2001
29:	894	0111	2002
30:	894	0111	2003
31:	894	0111	2004
32:	894	0111	2005
33:	894	0111	2006
34:	894	0111	2007
35:	894	0111	2008
36:	894	0111	2009
37:	894	0111	2010
38:	894	0111	2011
39:	894	0111	2012
40:	894	0111	2013
41:	804	0111	1994
42:	804	0111	1995
43:	804	0111	1996
44:	804	0111	1997
45:	804	0111	1998
46:	804	0111	1999
47:	804	0111	2000
48:	804	0111	2001
49:	804	0111	2002
50:	804	0111	2003
51:	804	0111	2004
52:	804	0111	2005
53:	804	0111	2006
54:	804	0111	2007
55:	804	0111	2008
56:	804	0111	2009
57:	804	0111	2010
58:	804	0111	2011
59:	804	0111	2012
60:	804	0111	2013
61:	400	0111	1994
62:	400	0111	1995
63:	400	0111	1996
64:	400	0111	1997
65:	400	0111	1998
66:	400	0111	1999

67:	400	0111	2000
68:	400	0111	2001
69:	400	0111	2002
70:	400	0111	2003
71:	400	0111	2004
72:	400	0111	2005
73:	400	0111	2006
74:	400	0111	2007
75:	400	0111	2008
76:	400	0111	2009
77:	400	0111	2010
78:	400	0111	2011
79:	400	0111	2012
80:	400	0111	2013
81:	348	0111	1994
82:	348	0111	1995
83:	348	0111	1996
84:	348	0111	1997
85:	348	0111	1998
86:	348	0111	1999
87:	348	0111	2000
88:	348	0111	2001
89:	348	0111	2002
90:	348	0111	2003
91:	348	0111	2004
92:	348	0111	2005
93:	348	0111	2006
94:	348	0111	2007
95:	348	0111	2008
96:	348	0111	2009
97:	348	0111	2010
98:	348	0111	2011
99:	348	0111	2012
100:	348	0111	2013

geographicAreaM49 measuredItemCPC timePointYears

Value\_measuredElement\_5212 flagObservationStatus\_measuredElement\_5212

1:	1319760	I
2:	1181120	I
3:	957670	I
4:	1211720	I
5:	1141682	I
6:	966282	I
7:	978575	I
8:	1355500	I
9:	1368627	I
10:	841014	I
11:	1039680	I
12:	1101807	I
13:	970392	I
14:	1087996	I
15:	1111533	I

16:	1247718	I
17:	1137650	I
18:	1137642	I
19:	1090000	I
20:	NA	M
21:	11566	I
22:	7806	I
23:	10327	I
24:	10693	I
25:	11278	I
26:	9921	I
27:	12077	I
28:	12000	I
29:	12000	I
30:	21000	I
31:	13543	I
32:	22033	I
33:	17100	I
34:	18833	I
35:	11394	I
36:	34296	I
37:	27291	I
38:	37631	I
39:	NA	M
40:	NA	M
41:	4507000	I
42:	5479400	I
43:	5891800	I
44:	6508400	I
45:	5641000	I
46:	5931600	I
47:	5161600	I
48:	6882000	I
49:	6749600	I
50:	2456400	I
51:	5533700	I
52:	6571000	I
53:	5511000	I
54:	5951300	I
55:	7053600	I
56:	6752900	I
57:	6284100	I
58:	6657300	I
59:	6400000	I
60:	NA	M
61:	29715	I
62:	40555	I
63:	28349	I
64:	37920	I
65:	28835	I
66:	4087	I

67:	18204	I
68:	13449	I
69:	32745	I
70:	29781	I
71:	10728	I
72:	29292	I
73:	26920	I
74:	20760	I
75:	12456	I
76:	15878	I
77:	21466	I
78:	14330	I
79:	NA	M
80:	NA	M
81:	1058749	I
82:	1108000	I
83:	1193340	I
84:	1247569	I
85:	1183540	I
86:	734100	I
87:	1024430	I
88:	1205610	I
89:	1110471	I
90:	1113755	I
91:	1173800	I
92:	1130719	I
93:	1074735	I
94:	1111269	I
95:	1125629	I
96:	1146456	I
97:	1011182	I
98:	978000	I
99:	1080000	I
100:	NA	M
Value_measuredElement_5212 flagObservationStatus_measuredElement_5212		
flagMethod_measuredElement_5212 Value_measuredElement_5312		
1:	e	1319760
2:	e	1181120
3:	e	957670
4:	e	1211720
5:	e	1141682
6:	e	966282
7:	e	978575
8:	e	1355500
9:	e	1368627
10:	e	841014
11:	e	1039680
12:	e	1101807
13:	e	970392
14:	e	1087996
15:	e	1111533

16:	e	1247718
17:	e	1137650
18:	e	1137642
19:	e	1090000
20:	e	NA
21:	e	11566
22:	e	7806
23:	e	10327
24:	e	10693
25:	e	11278
26:	e	9921
27:	e	12077
28:	e	12000
29:	e	12000
30:	e	21000
31:	e	13543
32:	e	22033
33:	e	17100
34:	e	18833
35:	e	11394
36:	e	34296
37:	e	27291
38:	e	37631
39:	e	NA
40:	e	NA
41:	e	4507000
42:	e	5479400
43:	e	5891800
44:	e	6508400
45:	e	5641000
46:	e	5931600
47:	e	5161600
48:	e	6882000
49:	e	6749600
50:	e	2456400
51:	e	5533700
52:	e	6571000
53:	e	5511000
54:	e	5951300
55:	e	7053600
56:	e	6752900
57:	e	6284100
58:	e	6657300
59:	e	6400000
60:	e	NA
61:	e	29715
62:	e	40555
63:	e	28349
64:	e	37920
65:	e	28835
66:	e	4087

67:	e	18204
68:	e	13449
69:	e	32745
70:	e	29781
71:	e	10728
72:	e	29292
73:	e	26920
74:	e	20760
75:	e	12456
76:	e	15878
77:	e	21466
78:	e	14330
79:	e	NA
80:	e	NA
81:	e	1058749
82:	e	1108000
83:	e	1193340
84:	e	1247569
85:	e	1183540
86:	e	734100
87:	e	1024430
88:	e	1205610
89:	e	1110471
90:	e	1113755
91:	e	1173800
92:	e	1130719
93:	e	1074735
94:	e	1111269
95:	e	1125629
96:	e	1146456
97:	e	1011182
98:	e	978000
99:	e	1080000
100:	e	NA
flagMethod_measuredElement_5212 Value_measuredElement_5312		
flagObservationStatus_measuredElement_5312 flagMethod_measuredElement_5312		
1:		-
2:		-
3:		-
4:		-
5:		-
6:		-
7:		-
8:		-
9:		-
10:		-
11:		-
12:		-
13:		-
14:		-
15:		-

16: -  
 17: -  
 18: -  
 19: T p  
 20: M u  
 21: -  
 22: -  
 23: -  
 24: -  
 25: -  
 26: -  
 27: -  
 28: T p  
 29: T p  
 30: E f  
 31: T p  
 32: -  
 33: T p  
 34: -  
 35: -  
 36: -  
 37: -  
 38: -  
 39: M u  
 40: M u  
 41: -  
 42: -  
 43: -  
 44: -  
 45: -  
 46: -  
 47: -  
 48: -  
 49: -  
 50: -  
 51: -  
 52: -  
 53: -  
 54: -  
 55: -  
 56: -  
 57: -  
 58: -  
 59: T p  
 60: M u  
 61: -  
 62: -  
 63: -  
 64: -  
 65: -  
 66: -

67:			-
68:			-
69:			-
70:			-
71:			-
72:			-
73:			-
74:			-
75:			-
76:			-
77:			-
78:			-
79:		M	u
80:		M	u
81:			-
82:			-
83:			-
84:			-
85:			-
86:			-
87:			-
88:			-
89:			-
90:			-
91:			-
92:			-
93:			-
94:			-
95:			-
96:			-
97:			-
98:			-
99:		T	p
100:		M	u
	flagObservationStatus_measuredElement_5312 flagMethod_measuredElement_5312		
	Value_measuredElement_5525 flagObservationStatus_measuredElement_5525		
1:	NA		M
2:	NA		M
3:	NA		M
4:	NA		M
5:	NA		M
6:	NA		M
7:	NA		M
8:	NA		M
9:	NA		M
10:	NA		M
11:	NA		M
12:	NA		M
13:	NA		M
14:	NA		M
15:	396000		



16:	753000	
17:	NA	M
18:	NA	M
19:	NA	M
20:	NA	M
21:	NA	M
22:	NA	M
23:	NA	M
24:	NA	M
25:	NA	M
26:	NA	M
27:	NA	M
28:	NA	M
29:	NA	M
30:	NA	M
31:	NA	M
32:	NA	M
33:	NA	M
34:	NA	M
35:	NA	M
36:	NA	M
37:	NA	M
38:	NA	M
39:	NA	M
40:	NA	M
41:	1450000	T
42:	1450000	T
43:	1450000	T
44:	1450000	T
45:	1450000	T
46:	1450000	T
47:	1450000	T
48:	1450000	T
49:	1450000	T
50:	700000	T
51:	1450000	T
52:	1540000	T
53:	1300000	T
54:	1500000	T
55:	1800000	T
56:	1750000	T
57:	1840000	T
58:	1800000	E
59:	NA	M
60:	NA	M
61:	7600	
62:	4500	
63:	5000	
64:	4500	
65:	5046	
66:	5066	

67:	4763	
68:	4400	
69:	4697	
70:	5507	
71:	4143	
72:	3665	
73:	4460	
74:	2097	
75:	2300	
76:	2325	
77:	2689	
78:	NA	M
79:	NA	M
80:	NA	M
81:	266220	
82:	275650	
83:	242885	
84:	322661	
85:	229715	
86:	214647	
87:	224580	
88:	239291	
89:	142524	
90:	56871	
91:	62857	
92:	282680	
93:	281626	
94:	281626	
95:	289033	
96:	278308	
97:	258829	
98:	260000	E
99:	NA	M
100:	NA	M
Value_measuredElement_5525 flagObservationStatus_measuredElement_5525		
flagMethod_measuredElement_5525 Value_areaSownRatio		
1:	u	1
2:	u	1
3:	u	1
4:	u	1
5:	u	1
6:	u	1
7:	u	1
8:	u	1
9:	u	1
10:	u	1
11:	u	1
12:	u	1
13:	u	1
14:	u	1
15:	-	1

16:	-	1
17:	u	1
18:	u	1
19:	u	1
20:	u	1
21:	u	1
22:	u	1
23:	u	1
24:	u	1
25:	u	1
26:	u	1
27:	u	1
28:	u	1
29:	u	1
30:	u	1
31:	u	1
32:	u	1
33:	u	1
34:	u	1
35:	u	1
36:	u	1
37:	u	1
38:	u	1
39:	u	1
40:	u	1
41:	p	1
42:	p	1
43:	p	1
44:	p	1
45:	p	1
46:	p	1
47:	p	1
48:	p	1
49:	p	1
50:	p	1
51:	p	1
52:	p	1
53:	p	1
54:	p	1
55:	p	1
56:	p	1
57:	p	1
58:	f	1
59:	u	1
60:	u	1
61:	-	1
62:	-	1
63:	-	1
64:	-	1
65:	-	1
66:	-	1

67:	-	1
68:	-	1
69:	-	1
70:	-	1
71:	-	1
72:	-	1
73:	-	1
74:	-	1
75:	-	1
76:	-	1
77:	-	1
78:	u	1
79:	u	1
80:	u	1
81:	-	1
82:	-	1
83:	-	1
84:	-	1
85:	-	1
86:	-	1
87:	-	1
88:	-	1
89:	-	1
90:	-	1
91:	-	1
92:	-	1
93:	-	1
94:	-	1
95:	-	1
96:	-	1
97:	-	1
98:	f	1
99:	u	1
100:	u	1

flagMethod\_measuredElement\_5525 Value\_areaSownRatio  
 flagObservationStatus\_areaSownRatio flagMethod\_areaSownRatio

1:	I	e
2:	I	e
3:	I	e
4:	I	e
5:	I	e
6:	I	e
7:	I	e
8:	I	e
9:	I	e
10:	I	e
11:	I	e
12:	I	e
13:	I	e
14:	I	e
15:	I	e

16: I e  
17: I e  
18: I e  
19: I e  
20: I e  
21: I e  
22: I e  
23: I e  
24: I e  
25: I e  
26: I e  
27: I e  
28: I e  
29: I e  
30: I e  
31: I e  
32: I e  
33: I e  
34: I e  
35: I e  
36: I e  
37: I e  
38: I e  
39: I e  
40: I e  
41: I e  
42: I e  
43: I e  
44: I e  
45: I e  
46: I e  
47: I e  
48: I e  
49: I e  
50: I e  
51: I e  
52: I e  
53: I e  
54: I e  
55: I e  
56: I e  
57: I e  
58: I e  
59: I e  
60: I e  
61: I e  
62: I e  
63: I e  
64: I e  
65: I e  
66: I e

```

67:          I          e
68:          I          e
69:          I          e
70:          I          e
71:          I          e
72:          I          e
73:          I          e
74:          I          e
75:          I          e
76:          I          e
77:          I          e
78:          I          e
79:          I          e
80:          I          e
81:          I          e
82:          I          e
83:          I          e
84:          I          e
85:          I          e
86:          I          e
87:          I          e
88:          I          e
89:          I          e
90:          I          e
91:          I          e
92:          I          e
93:          I          e
94:          I          e
95:          I          e
96:          I          e
97:          I          e
98:          I          e
99:          I          e
100:         I          e
      flagObservationStatus_areaSownRatio flagMethod_areaSownRatio

```

```

> data[geographicAreaM49 == 100,
+      .(time = timePointYears, areaSown = Value_measuredElement_5212,
+      areaHarvested = Value_measuredElement_5312, Value_areaSownRatio)]

```

	time	areaSown	areaHarvested	Value_areaSownRatio
1:	1994	1319760	1319760	1
2:	1995	1181120	1181120	1
3:	1996	957670	957670	1
4:	1997	1211720	1211720	1
5:	1998	1141682	1141682	1
6:	1999	966282	966282	1
7:	2000	978575	978575	1
8:	2001	1355500	1355500	1
9:	2002	1368627	1368627	1
10:	2003	841014	841014	1
11:	2004	1039680	1039680	1

12: 2005	1101807	1101807	1
13: 2006	970392	970392	1
14: 2007	1087996	1087996	1
15: 2008	1111533	1111533	1
16: 2009	1247718	1247718	1
17: 2010	1137650	1137650	1
18: 2011	1137642	1137642	1
19: 2012	1090000	1090000	1
20: 2013	NA	NA	1

```
> imputeAreaSown(data = data, codeAreaSown = "5212")
```

	geographicAreaM49	measuredItemCPC	timePointYears
1:	100	0111	1994
2:	100	0111	1995
3:	100	0111	1996
4:	100	0111	1997
5:	100	0111	1998
6:	100	0111	1999
7:	100	0111	2000
8:	100	0111	2001
9:	100	0111	2002
10:	100	0111	2003
11:	100	0111	2004
12:	100	0111	2005
13:	100	0111	2006
14:	100	0111	2007
15:	100	0111	2008
16:	100	0111	2009
17:	100	0111	2010
18:	100	0111	2011
19:	100	0111	2012
20:	100	0111	2013
21:	894	0111	1994
22:	894	0111	1995
23:	894	0111	1996
24:	894	0111	1997
25:	894	0111	1998
26:	894	0111	1999
27:	894	0111	2000
28:	894	0111	2001
29:	894	0111	2002
30:	894	0111	2003
31:	894	0111	2004
32:	894	0111	2005
33:	894	0111	2006
34:	894	0111	2007
35:	894	0111	2008
36:	894	0111	2009
37:	894	0111	2010
38:	894	0111	2011
39:	894	0111	2012

40:	894	0111	2013
41:	804	0111	1994
42:	804	0111	1995
43:	804	0111	1996
44:	804	0111	1997
45:	804	0111	1998
46:	804	0111	1999
47:	804	0111	2000
48:	804	0111	2001
49:	804	0111	2002
50:	804	0111	2003
51:	804	0111	2004
52:	804	0111	2005
53:	804	0111	2006
54:	804	0111	2007
55:	804	0111	2008
56:	804	0111	2009
57:	804	0111	2010
58:	804	0111	2011
59:	804	0111	2012
60:	804	0111	2013
61:	400	0111	1994
62:	400	0111	1995
63:	400	0111	1996
64:	400	0111	1997
65:	400	0111	1998
66:	400	0111	1999
67:	400	0111	2000
68:	400	0111	2001
69:	400	0111	2002
70:	400	0111	2003
71:	400	0111	2004
72:	400	0111	2005
73:	400	0111	2006
74:	400	0111	2007
75:	400	0111	2008
76:	400	0111	2009
77:	400	0111	2010
78:	400	0111	2011
79:	400	0111	2012
80:	400	0111	2013
81:	348	0111	1994
82:	348	0111	1995
83:	348	0111	1996
84:	348	0111	1997
85:	348	0111	1998
86:	348	0111	1999
87:	348	0111	2000
88:	348	0111	2001
89:	348	0111	2002
90:	348	0111	2003



91:	348	0111	2004
92:	348	0111	2005
93:	348	0111	2006
94:	348	0111	2007
95:	348	0111	2008
96:	348	0111	2009
97:	348	0111	2010
98:	348	0111	2011
99:	348	0111	2012
100:	348	0111	2013

geographicAreaM49 measuredItemCPC timePointYears

	Value_measuredElement_5212	flagObservationStatus_measuredElement_5212
1:	1319760	I
2:	1181120	I
3:	957670	I
4:	1211720	I
5:	1141682	I
6:	966282	I
7:	978575	I
8:	1355500	I
9:	1368627	I
10:	841014	I
11:	1039680	I
12:	1101807	I
13:	970392	I
14:	1087996	I
15:	1111533	I
16:	1247718	I
17:	1137650	I
18:	1137642	I
19:	1090000	I
20:	NA	M
21:	11566	I
22:	7806	I
23:	10327	I
24:	10693	I
25:	11278	I
26:	9921	I
27:	12077	I
28:	12000	I
29:	12000	I
30:	21000	I
31:	13543	I
32:	22033	I
33:	17100	I
34:	18833	I
35:	11394	I
36:	34296	I
37:	27291	I
38:	37631	I
39:	NA	M

40:	NA	M
41:	4507000	I
42:	5479400	I
43:	5891800	I
44:	6508400	I
45:	5641000	I
46:	5931600	I
47:	5161600	I
48:	6882000	I
49:	6749600	I
50:	2456400	I
51:	5533700	I
52:	6571000	I
53:	5511000	I
54:	5951300	I
55:	7053600	I
56:	6752900	I
57:	6284100	I
58:	6657300	I
59:	6400000	I
60:	NA	M
61:	29715	I
62:	40555	I
63:	28349	I
64:	37920	I
65:	28835	I
66:	4087	I
67:	18204	I
68:	13449	I
69:	32745	I
70:	29781	I
71:	10728	I
72:	29292	I
73:	26920	I
74:	20760	I
75:	12456	I
76:	15878	I
77:	21466	I
78:	14330	I
79:	NA	M
80:	NA	M
81:	1058749	I
82:	1108000	I
83:	1193340	I
84:	1247569	I
85:	1183540	I
86:	734100	I
87:	1024430	I
88:	1205610	I
89:	1110471	I
90:	1113755	I

91:	1173800	I
92:	1130719	I
93:	1074735	I
94:	1111269	I
95:	1125629	I
96:	1146456	I
97:	1011182	I
98:	978000	I
99:	1080000	I
100:	NA	M

Value\_measuredElement\_5212 flagObservationStatus\_measuredElement\_5212  
 flagMethod\_measuredElement\_5212 Value\_measuredElement\_5312

1:	e	1319760
2:	e	1181120
3:	e	957670
4:	e	1211720
5:	e	1141682
6:	e	966282
7:	e	978575
8:	e	1355500
9:	e	1368627
10:	e	841014
11:	e	1039680
12:	e	1101807
13:	e	970392
14:	e	1087996
15:	e	1111533
16:	e	1247718
17:	e	1137650
18:	e	1137642
19:	e	1090000
20:	e	NA
21:	e	11566
22:	e	7806
23:	e	10327
24:	e	10693
25:	e	11278
26:	e	9921
27:	e	12077
28:	e	12000
29:	e	12000
30:	e	21000
31:	e	13543
32:	e	22033
33:	e	17100
34:	e	18833
35:	e	11394
36:	e	34296
37:	e	27291
38:	e	37631
39:	e	NA

40:	e	NA
41:	e	4507000
42:	e	5479400
43:	e	5891800
44:	e	6508400
45:	e	5641000
46:	e	5931600
47:	e	5161600
48:	e	6882000
49:	e	6749600
50:	e	2456400
51:	e	5533700
52:	e	6571000
53:	e	5511000
54:	e	5951300
55:	e	7053600
56:	e	6752900
57:	e	6284100
58:	e	6657300
59:	e	6400000
60:	e	NA
61:	e	29715
62:	e	40555
63:	e	28349
64:	e	37920
65:	e	28835
66:	e	4087
67:	e	18204
68:	e	13449
69:	e	32745
70:	e	29781
71:	e	10728
72:	e	29292
73:	e	26920
74:	e	20760
75:	e	12456
76:	e	15878
77:	e	21466
78:	e	14330
79:	e	NA
80:	e	NA
81:	e	1058749
82:	e	1108000
83:	e	1193340
84:	e	1247569
85:	e	1183540
86:	e	734100
87:	e	1024430
88:	e	1205610
89:	e	1110471
90:	e	1113755

91:	e	1173800
92:	e	1130719
93:	e	1074735
94:	e	1111269
95:	e	1125629
96:	e	1146456
97:	e	1011182
98:	e	978000
99:	e	1080000
100:	e	NA

flagMethod\_measuredElement\_5212 Value\_measuredElement\_5312  
flagObservationStatus\_measuredElement\_5312 flagMethod\_measuredElement\_5312

1:		-
2:		-
3:		-
4:		-
5:		-
6:		-
7:		-
8:		-
9:		-
10:		-
11:		-
12:		-
13:		-
14:		-
15:		-
16:		-
17:		-
18:		-
19:	T	p
20:	M	u
21:		-
22:		-
23:		-
24:		-
25:		-
26:		-
27:		-
28:	T	p
29:	T	p
30:	E	f
31:	T	p
32:		-
33:	T	p
34:		-
35:		-
36:		-
37:		-
38:		-
39:	M	u

40:	M	u
41:		-
42:		-
43:		-
44:		-
45:		-
46:		-
47:		-
48:		-
49:		-
50:		-
51:		-
52:		-
53:		-
54:		-
55:		-
56:		-
57:		-
58:		-
59:	T	p
60:	M	u
61:		-
62:		-
63:		-
64:		-
65:		-
66:		-
67:		-
68:		-
69:		-
70:		-
71:		-
72:		-
73:		-
74:		-
75:		-
76:		-
77:		-
78:		-
79:	M	u
80:	M	u
81:		-
82:		-
83:		-
84:		-
85:		-
86:		-
87:		-
88:		-
89:		-
90:		-

```

91: -
92: -
93: -
94: -
95: -
96: -
97: -
98: -
99: T p
100: M u
    flagObservationStatus_measuredElement_5312 flagMethod_measuredElement_5312
    Value_measuredElement_5525 flagObservationStatus_measuredElement_5525
1: NA M
2: NA M
3: NA M
4: NA M
5: NA M
6: NA M
7: NA M
8: NA M
9: NA M
10: NA M
11: NA M
12: NA M
13: NA M
14: NA M
15: 396000
16: 753000
17: NA M
18: NA M
19: NA M
20: NA M
21: NA M
22: NA M
23: NA M
24: NA M
25: NA M
26: NA M
27: NA M
28: NA M
29: NA M
30: NA M
31: NA M
32: NA M
33: NA M
34: NA M
35: NA M
36: NA M
37: NA M
38: NA M
39: NA M

```

40:	NA	M
41:	1450000	T
42:	1450000	T
43:	1450000	T
44:	1450000	T
45:	1450000	T
46:	1450000	T
47:	1450000	T
48:	1450000	T
49:	1450000	T
50:	700000	T
51:	1450000	T
52:	1540000	T
53:	1300000	T
54:	1500000	T
55:	1800000	T
56:	1750000	T
57:	1840000	T
58:	1800000	E
59:	NA	M
60:	NA	M
61:	7600	
62:	4500	
63:	5000	
64:	4500	
65:	5046	
66:	5066	
67:	4763	
68:	4400	
69:	4697	
70:	5507	
71:	4143	
72:	3665	
73:	4460	
74:	2097	
75:	2300	
76:	2325	
77:	2689	
78:	NA	M
79:	NA	M
80:	NA	M
81:	266220	
82:	275650	
83:	242885	
84:	322661	
85:	229715	
86:	214647	
87:	224580	
88:	239291	
89:	142524	
90:	56871	



91:	62857	
92:	282680	
93:	281626	
94:	281626	
95:	289033	
96:	278308	
97:	258829	
98:	260000	E
99:	NA	M
100:	NA	M

Value_measuredElement_5525	flagObservationStatus_measuredElement_5525
flagMethod_measuredElement_5525	Value_areaSownRatio

1:	u	1
2:	u	1
3:	u	1
4:	u	1
5:	u	1
6:	u	1
7:	u	1
8:	u	1
9:	u	1
10:	u	1
11:	u	1
12:	u	1
13:	u	1
14:	u	1
15:	-	1
16:	-	1
17:	u	1
18:	u	1
19:	u	1
20:	u	1
21:	u	1
22:	u	1
23:	u	1
24:	u	1
25:	u	1
26:	u	1
27:	u	1
28:	u	1
29:	u	1
30:	u	1
31:	u	1
32:	u	1
33:	u	1
34:	u	1
35:	u	1
36:	u	1
37:	u	1
38:	u	1
39:	u	1

40:	u	1
41:	p	1
42:	p	1
43:	p	1
44:	p	1
45:	p	1
46:	p	1
47:	p	1
48:	p	1
49:	p	1
50:	p	1
51:	p	1
52:	p	1
53:	p	1
54:	p	1
55:	p	1
56:	p	1
57:	p	1
58:	f	1
59:	u	1
60:	u	1
61:	-	1
62:	-	1
63:	-	1
64:	-	1
65:	-	1
66:	-	1
67:	-	1
68:	-	1
69:	-	1
70:	-	1
71:	-	1
72:	-	1
73:	-	1
74:	-	1
75:	-	1
76:	-	1
77:	-	1
78:	u	1
79:	u	1
80:	u	1
81:	-	1
82:	-	1
83:	-	1
84:	-	1
85:	-	1
86:	-	1
87:	-	1
88:	-	1
89:	-	1
90:	-	1

91:	-	1
92:	-	1
93:	-	1
94:	-	1
95:	-	1
96:	-	1
97:	-	1
98:	f	1
99:	u	1
100:	u	1

  

flagMethod_measuredElement_5525 Value_areaSownRatio		
flagObservationStatus_areaSownRatio flagMethod_areaSownRatio		
1:	I	e
2:	I	e
3:	I	e
4:	I	e
5:	I	e
6:	I	e
7:	I	e
8:	I	e
9:	I	e
10:	I	e
11:	I	e
12:	I	e
13:	I	e
14:	I	e
15:	I	e
16:	I	e
17:	I	e
18:	I	e
19:	I	e
20:	I	e
21:	I	e
22:	I	e
23:	I	e
24:	I	e
25:	I	e
26:	I	e
27:	I	e
28:	I	e
29:	I	e
30:	I	e
31:	I	e
32:	I	e
33:	I	e
34:	I	e
35:	I	e
36:	I	e
37:	I	e
38:	I	e
39:	I	e

40:	I	e
41:	I	e
42:	I	e
43:	I	e
44:	I	e
45:	I	e
46:	I	e
47:	I	e
48:	I	e
49:	I	e
50:	I	e
51:	I	e
52:	I	e
53:	I	e
54:	I	e
55:	I	e
56:	I	e
57:	I	e
58:	I	e
59:	I	e
60:	I	e
61:	I	e
62:	I	e
63:	I	e
64:	I	e
65:	I	e
66:	I	e
67:	I	e
68:	I	e
69:	I	e
70:	I	e
71:	I	e
72:	I	e
73:	I	e
74:	I	e
75:	I	e
76:	I	e
77:	I	e
78:	I	e
79:	I	e
80:	I	e
81:	I	e
82:	I	e
83:	I	e
84:	I	e
85:	I	e
86:	I	e
87:	I	e
88:	I	e
89:	I	e
90:	I	e

```

91:                                I                                e
92:                                I                                e
93:                                I                                e
94:                                I                                e
95:                                I                                e
96:                                I                                e
97:                                I                                e
98:                                I                                e
99:                                I                                e
100:                               I                                e
      flagObservationStatus_areaSownRatio flagMethod_areaSownRatio

> data[geographicAreaM49 == 100,
+      .(time = timePointYears, areaSown = Value_measuredElement_5212,
+      areaHarvested = Value_measuredElement_5312, Value_areaSownRatio)]

      time areaSown areaHarvested Value_areaSownRatio
1: 1994   1319760      1319760             1
2: 1995   1181120      1181120             1
3: 1996    957670       957670             1
4: 1997   1211720      1211720             1
5: 1998   1141682      1141682             1
6: 1999    966282       966282             1
7: 2000    978575       978575             1
8: 2001   1355500      1355500             1
9: 2002   1368627      1368627             1
10: 2003    841014       841014             1
11: 2004   1039680      1039680             1
12: 2005   1101807      1101807             1
13: 2006    970392       970392             1
14: 2007   1087996      1087996             1
15: 2008   1111533      1111533             1
16: 2009   1247718      1247718             1
17: 2010   1137650      1137650             1
18: 2011   1137642      1137642             1
19: 2012   1090000      1090000             1
20: 2013         NA         NA             1

```

The two cases above are identical. Using an ensemble provides no advantage as no area sown data is originally available.

### 2.3. Estimate Seed Rate

For this vignette, we will proceed with the dataset defined above: seedData with a CPC code of 0111.

The next step in getting the seed usage is to estimate the seeding rate. The database stores two tables, default\_seed\_rate and specific\_seed\_rate, which contain estimates for seed rates. The specific seed rate table contains values for country commodity pairs, while the default seed rate contains average values for commodities overall. It would therefore be preferable to use the specific\_seed\_rate table, but values are not always available for all countries. Thus, the default\_seed\_rate table is used when entries are not available in the specific\_seed\_rate table.

```

> # countrySpecificData = getCountrySpecificSeedRate()
> countrySpecificData = data.table(
+   geographicAreaM49 = c("100", "348", "400"),
+   measuredItemCPC = "0111",
+   Value_seedRate = c(222, 213, 115),
+   flagObservationStatus_seedRate = c("E", "E", ""))
> setkeyv(countrySpecificData, c("geographicAreaM49", "measuredItemCPC"))
> fillCountrySpecificSeedRate(data = data,
+                               countrySpecificData = countrySpecificData)
> head(data, 1)

   geographicAreaM49 measuredItemCPC timePointYears Value_measuredElement_5212
1:                100             0111          1994             1319760
   flagObservationStatus_measuredElement_5212 flagMethod_measuredElement_5212
1:                                           I                             e
   Value_measuredElement_5312 flagObservationStatus_measuredElement_5312
1:             1319760
   flagMethod_measuredElement_5312 Value_measuredElement_5525
1:                -                             NA
   flagObservationStatus_measuredElement_5525 flagMethod_measuredElement_5525
1:                                           M                             u
   Value_areaSownRatio flagObservationStatus_areaSownRatio
1:                1                             I
   flagMethod_areaSownRatio Value_seedRate flagObservationStatus_seedRate
1:                e             222                             E

> data[, Value_seedRate]

 [1] 222 222 222 222 222 222 222 222 222 222 222 222 222 222 222 222 222 222
[19] 222 222 213 213 213 213 213 213 213 213 213 213 213 213 213 213 213 213
[37] 213 213 213 213 115 115 115 115 115 115 115 115 115 115 115 115 115 115
[55] 115 115 115 115 115 115  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA
[73]  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA
[91]  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA  NA

```

The `getCountrySpecificSeedRate` function simply pulls the `specific_seed_rate` table from the database. This table is also the default value for `countrySpecificData` in the `fillCountrySpecificSeedRate` function, so generally it will not need to be created as here (although it could be created manually from the commented out line above). However, for vignette creation, it was simpler to just create the part of the table we used at the time of the writing, and this is what is done here. The `fillCountrySpecificSeedRate` function adds an additional two columns to data with the `seedRate` value and observation flags.

```

> # generalSeedData = getCountryGeneralSeedRate()
> generalSeedData = data.table(measuredItemCPC = "0111", Value_seedRate = 151.14,
+                               flagObservationStatus_seedRate = "")
> setkeyv(generalSeedData, "measuredItemCPC")
> fillGeneralSeedRate(data = data,
+                       generalSeedData = generalSeedData)
> head(data, 1)

```

```

geographicAreaM49 measuredItemCPC timePointYears Value_measuredElement_5212
1:          100          0111          1994          1319760
flagObservationStatus_measuredElement_5212 flagMethod_measuredElement_5212
1:                                     I                                     e
Value_measuredElement_5312 flagObservationStatus_measuredElement_5312
1:          1319760
flagMethod_measuredElement_5312 Value_measuredElement_5525
1:          -          NA
flagObservationStatus_measuredElement_5525 flagMethod_measuredElement_5525
1:                                     M                                     u
Value_areaSownRatio flagObservationStatus_areaSownRatio
1:          1          I
flagMethod_areaSownRatio Value_seedRate flagObservationStatus_seedRate
1:          e          222          E

> data[, Value_seedRate]

```

```

[1] 222.00 222.00 222.00 222.00 222.00 222.00 222.00 222.00 222.00 222.00
[11] 222.00 222.00 222.00 222.00 222.00 222.00 222.00 222.00 222.00 222.00
[21] 213.00 213.00 213.00 213.00 213.00 213.00 213.00 213.00 213.00 213.00
[31] 213.00 213.00 213.00 213.00 213.00 213.00 213.00 213.00 213.00 213.00
[41] 115.00 115.00 115.00 115.00 115.00 115.00 115.00 115.00 115.00 115.00
[51] 115.00 115.00 115.00 115.00 115.00 115.00 115.00 115.00 115.00 115.00
[61] 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14
[71] 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14
[81] 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14
[91] 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14 151.14

```

This function updates all the NA values to the (“global”) default seeding rate for this commodity.

## 2.4. Estimate Seed Usage

The seed usage is estimated by way of the `imputeSeed` function:

Not all seeds have been imputed. To understand why, note that area sown is available for all years except the last year. But, area sown on year  $t$  corresponds to seed usage on year  $t - 1$ , thus we don’t have enough valid observations to impute either of the last two seed usages. These values remain missing, as seen above.

## 2.5. Push data back to database

Lastly, we must push the imputed data back to the database. This is done via the `saveSeedData` function (not evaluated, as the token would need to be valid).

```
> saveSeedData(data)
```

Note: this function assumes certain `swsContext` files exist within your workspace, although it does not require you to pass them to it. So, be careful not to rm those files!

## Affiliation:

Joshua M. Browning  
Economics and Social Statistics Division (ESS)

Economic and Social Development Department (ES)  
Food and Agriculture Organization of the United Nations (FAO)  
Viale delle Terme di Caracalla 00153 Rome, Italy  
E-mail: [joshua.browning@fao.org](mailto:joshua.browning@fao.org)  
URL: [https://github.com/rockclimber112358/sws\\_seed](https://github.com/rockclimber112358/sws_seed)

DRAFT