Analysis of Experimental Design for Phenotyping by Drought and Heat Resistance.

Juan Burgueño*[1]; Martha C. Willcox[1]; Armando E. Guadarrama[1]; Daniel Chepetla[1]; Enrique Rodríguez[1]; Noel Orlando Gómez Montiel[2]; Alejandro Ortega y Corona[2]; Ivan Ortiz-Monestario[1]; Sam Trachsel[1]; Edgar Antonio[1]; Armando Espinoza Banda[3]; Peter Wenzl[1]; Sarah Hearne[1].

1. CIMMYT, Int. México. 2. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias; 3. Universidad Autónoma Agraria Antonio Narro. * jburgueno@cgiar.org

INTRODUCTION Genetic resources have historically been characterized by taking data in the field, but recently, genotyping has become cheaper and easier than phenotyping. This change has necessitated linkage of genotypic and phenotypic experimental design and analysis strategies, with designs focusing on evaluation of large numbers of genotypes. The Seeds of Discovery Initiative has as one of its priorities the identification of germplasm and alleles to address climate change. In this paper we present the strategy we followed for phenotyping characterization of the testcrosses of the CIMMYT maize Genebank Breeder's Core Collection for heat and drought tolerance.

EXPERIMENTAL DESIGN AND MODEL ANALYSIS

Range																				
	border crop																			
20	400	399	398	397	396	395	394	393	392	391	390	389	388	387	386	385	384	383	382	381
	road																			
19	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380
18	360	359	358	357	356	355	354	353	352	351	350	349	348	347	346	345	344	343	342	341
17	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340
16	320	319	318	317	316	315	314	313	312	311	310	309	308	307	306	305	304	303	302	301
15	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300
14	280	279	278	277	276	275	274	273	272	271	270	269	268	267	266	265	264	263	262	261
13	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260
12	240	239	238	237	236	235	234	233	232	231	230	229	228	227	226	225	224	223	222	221
11	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220
10	200	199	198	197	196	195	194	193	192	191	190	189	188	187	186	185	184	183	182	181
9	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
8	160	159	158	157	156	155	154	153	152	151	150	149	148	147	146	145	144	143	142	141
7	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140
6	120	119	118	117	116	115	114	113	112	111	110	109	108	107	106	105	104	103	102	101
5	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
4	80	79	78	77	76	75	74	73	72	71	70	69	68	67	66	65	64	63	62	61
3	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
2	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	border crop																			
Row	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

Commercial hybrid 1 — Resistant check — Number of plots: 400

Commercial hybrid 2 — Susceptible check — 342 crosses / 58 checks

Augmented row column design

The same experimental design for irrigated (**WW**) and water stressed (**WS**)

Univariate model

$$Y_{lm} = \mu + T_i + P(T)_{ji} + A(PT)_{kji} + Row_l + Col_m + \varepsilon_{lm}$$

T: Checks j=1,2,...,t (4) (1=crosses)

P: Testers k=1,2,...,p (4)

A: Accesions (random) l=1,2,...,a (432)

$$\varepsilon \sim N\left(0, \Sigma = AR1_{row} \otimes AR1_{col}\right)$$

Multivariate model

$$Y_{plm} = \mu + M_p + T_i + MT_{pi} + P(T)_{ji} + MP(T)_{pji} + A(PT)_{kji} + MA(PT)_{pkji} + Row_l + Col_m + \varepsilon_{plm}$$

T: Checks j=1,2,...,t (4) (1=crosses)

P: Testers k=1,2,...,p (4)

A: Accesions (random) l=1,2,...,a (432)

M: Management p=1 (irrigated), 2 (water stressed)

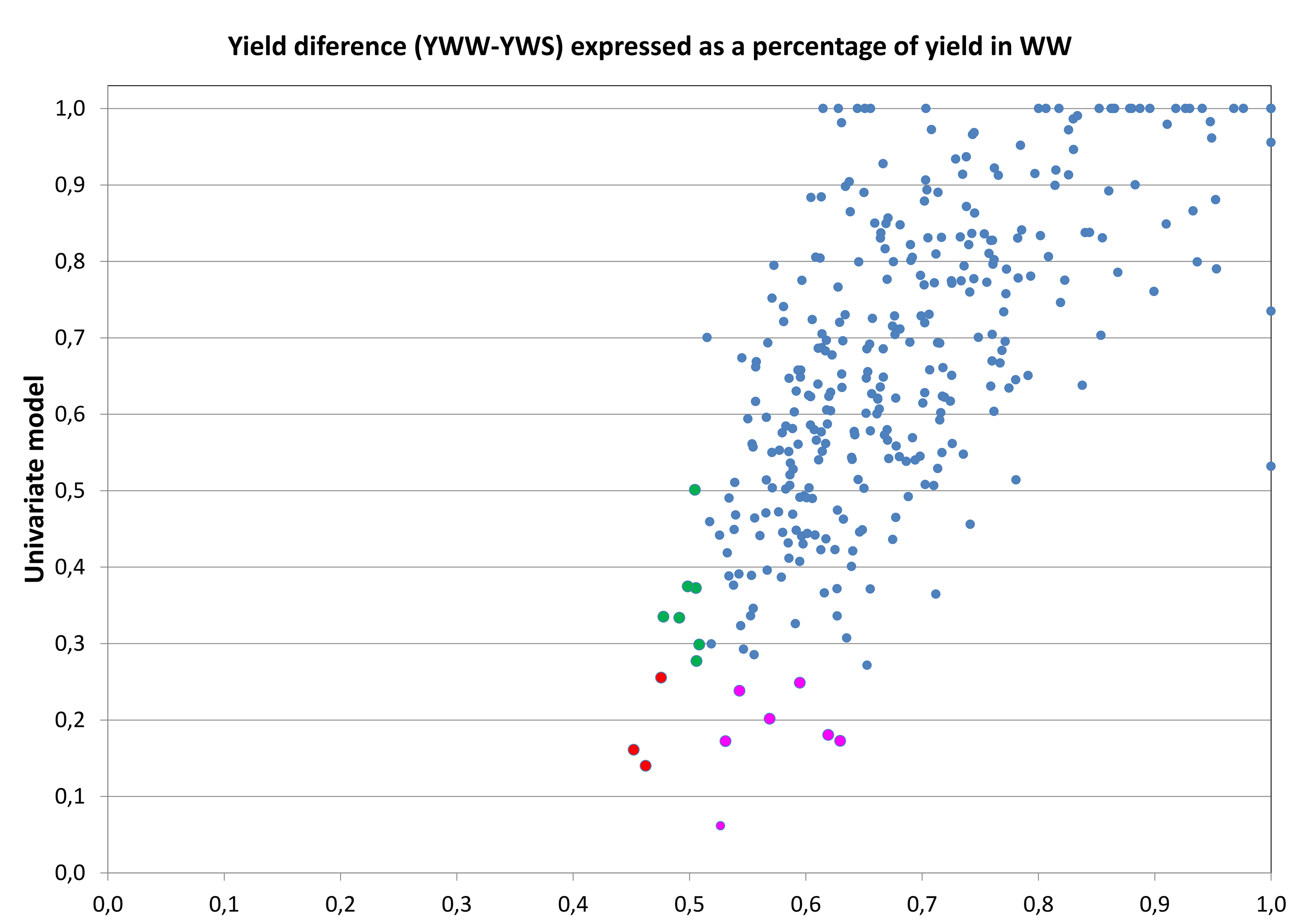
$$\varepsilon_p \sim N\left(0, \Sigma = AR1_{row} \otimes AR1_{col}\right)$$

$$cor(\varepsilon_1 \varepsilon_2) = \rho$$

	WW mv	WS mv	WW ind	WS ind
Std Err of Difference for Yield	0.846	0.846	0.945	1.942

Correlation between results obtained with the different models

		Observed				Individual analysis				Multivariate analysis		
		WW	WS	Diff	Diff %	WW	WS	Diff	Diff %	WW	WS	Diff
Observed	WS	0.340										
	Diff	0.679	-0.459									
	Diff %	-0.004	-0.890	0.691								
Individual analysis	WW	0.802	0.205	0.598	0.048							
	WS	0.202	0.430	-0.144	-0.378	0.262						
	Diff	0.079	-0.370	0.364	0.407	0.089	-0.938					
	Diff %	-0.126	-0.415	0.205	0.395	-0.164	-0.992	0.965				
Multivariate analysis	WW	0.607	0.227	0.395	-0.042	0.695	0.692	-0.465	-0.638			
	WS	0.414	0.230	0.211	-0.112	0.475	0.745	-0.598	-0.714	0.940		
	Diff	0.463	-0.048	0.475	0.214	0.530	-0.273	0.472	0.332	0.005	-0.335	
	Diff %	-0.351	-0.214	-0.164	0.117	-0.388	-0.721	0.605	0.706	-0.894	-0.980	0.405



CONCLUSIONS From a statistical point of view, multivariate analysis is more efficient because it incorporates all information and the correlation between environments and traits can be modeled. In the two step analysis or univariate model some information is lost compared to the one step analysis in which information from several sources is combined and improves predictions.