

Report on the Systematic Reconnaissance Flight Wading Bird Survey in Everglades National Park, 1985–199~~6~~⁸

Gareth J. Russell¹, Kenneth M. Portier², Oron. L. Bass³, D. Martin Fleming⁴, Larry D. Harris⁵, Christine Steible², Laurie Oberhofer³, and Stuart L. Pimm^{1*}

¹ Department of Ecology and Evolutionary Biology, University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996–1610.

² Department of Statistics, University of Florida, Gainesville, FL 32611.

³ South Florida Natural Resources Center, Everglades National Park, State Road ?????, Homestead, FL 33034.

⁴ PO Box 140703, Gainesville, FL 32614.

⁵ Program in Landscape Ecology and Biodiversity Conservation, Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611.

DRAFT COPY

1 Table of Contents

1	TABLE OF CONTENTS	2
2	EXECUTIVE SUMMARY	6
3	PREFACE	6
3.1	A note on the ordering of Figures and Tables	7
4	INTRODUCTION TO THE EVERGLADES	7
4.1	Location of the Everglades	7
4.2	Location of the Study Site	7
4.3	Climate	8
4.4	Physiography and geology	8
4.5	Soils	10
4.6	Vegetation	10
4.7	Animals	11
4.8	Hydrological history of the Everglades Landscape	11
4.8.1	Pre-drainage	11
4.8.2	Post-drainage: General	13
4.8.3	Post-drainage: Freshwater wetland hydrodynamics	14
4.8.4	Post-drainage: Coastal (estuarine) wetland hydrodynamics	14
4.8.5	Overview	14
4.9	Related changes in fundamental landscape characteristics	15
4.10	Animal issues of concern	15
4.11	Historic and present status of colonial wading bird populations	16
4.12	Project objectives (management questions)	17
5	THE SRF SURVEY	18
5.1	The SRF approach	18
5.2	History of the SRF project	18
5.3	The SRF protocol	18
5.4	Mapping surface water patterns	19
5.5	Other variables of interest	20
5.6	Notation	20
5.7	Precision and Accuracy of Counts	21
5.8	Precision	21
5.8.1	Method	21
5.8.2	Variance of the Total Density	22
5.8.3	Variance of Transect Density	23
5.8.4	Variance of the Individual Cells	24
5.8.5	Alternative Cell Variance Estimation Procedures	25
5.8.6	Summary	27
5.8.7	Future Work	27
5.9	Accuracy	27
5.9.1	Counting bias	28
5.9.2	Methods	28
5.9.3	Analysis of Variance	29
5.9.4	Summary (ANOVA)	29
5.9.5	Regression	30
5.9.6	Summary	30
5.9.7	Future Work	31

DRAFT COPY

5.10	Spatial Distribution and Grid Cell Resolution	31
5.10.1	<i>Methods</i>	31
5.10.2	<i>Explanation of the Resampling Scheme</i>	32
5.10.3	<i>Statistics Examined</i>	32
5.10.4	<i>Results</i>	32
5.10.5	<i>Future Work</i>	33
6	SURFACE WATER CONDITION: TEMPORAL PATTERNS	34
7	BIRD COUNTS: TEMPORAL PATTERNS	35
7.1	Temporal autocorrelation	35
7.1.1	<i>Summary (Total Birds)</i>	35
7.1.2	<i>Summary (By Size and Color Class)</i>	35
7.1.3	<i>Summary (Size and Color Class by Basin)</i>	36
7.1.4	<i>Future Work</i>	36
8	BIRD COUNTS: SPATIAL PATTERNS	36
8.1	Spatial autocorrelation within species	36
8.2	Associations Among Guild/Species	36
8.3	Spatial patterns 1: by basin	36
8.4	Spatial patterns 2: by cell	36
9	BIRD DENSITY ESTIMATION	37
10	BIRD DENSITY PREDICTION: LINEAR MODELS AND LAND ELEMENT THEMES	38
10.1	Decomposition of total variability in bird counts and importance of environmental correlates	39
10.2	<i>Methods</i>	40
10.3	Identification of land element themes	41
10.3.1	<i>Dominant elevations</i>	41
10.3.2	<i>Dominant ground surface features</i>	41
10.3.3	<i>Dominant soil types</i>	41
10.3.4	<i>Dominant vegetation cover types</i>	41
10.3.5	<i>Surface water condition</i>	42
10.3.6	<i>Surface water hydroperiods</i>	42
10.3.7	<i>Landscape classification</i>	42
10.4	<i>Results</i>	42
10.4.1	<i>Temporal Trend</i>	42
10.4.2	<i>Spatial Pattern</i>	43
10.4.3	<i>Correlations to Land Elements</i>	43
10.4.4	<i>Correlations to Landscape Classifications</i>	43
10.4.5	<i>Correlations to Surface Water Condition</i>	43
10.4.6	<i>Tree Regression</i>	43
10.5	<i>Summary of linear modelling</i>	43
11	BIRD DENSITY PREDICTION: A SPATIAL MODEL OF WADING BIRD DISTRIBUTION AND ABUNDANCE IN EVERGLADES NATIONAL PARK	44
11.1	<i>Abstract</i>	44
12	OVERALL CONCLUSIONS AND RECOMMENDATIONS	44
12.1	<i>The future of the SRF</i>	44
12.2	<i>Possible changes in the SRF</i>	45
12.3	<i>Looking ahead: including nesting</i>	45
13	ACKNOWLEDGMENTS	45
14	REFERENCES	45
15	APPENDIX A: LINEAR MODEL METHODOLOGY	48
15.1	<i>Generalized Linear Model</i>	48
15.2	<i>Generalized Linear Mixed Model</i>	50

16	APPENDIX B: STATISTICAL CLUSTER CLASSIFICATION DEVELOPMENT.....	52
17	APPENDIX C: COMPUTING DRYING (P_D) AND WETTING (P_W) PROPORTIONS AND THE CHANGE POTENTIAL INDEX (CPI).....	53
18	APPENDIX D: TREE REGRESSION METHODOLOGY.....	53
19	FIGURES.....	56
19.1	Figure 1. Map of southern Everglades and Big Cypress National Preserve with sampling grid and flight line	56
19.2	Figure 2. SRF sampling protocol. (Figure of fixed wing aircraft with strip width, height, speed).....	57
19.3	Figure 3. Dominant elevations in the SRF study area.....	58
19.4	Figure 4. Dominant soil types in the SRF study area.....	59
19.5	Figure 5. Dominant vegetation in the SRF study area	60
19.6	Figure 6. Basin regions for the SRF study area.....	61
19.7	Figure 7. Physiographic regions for the SRF study area	62
19.8	Figure 8. Watershed regions for the SRF study area.....	63
19.9	Figure 9. Pond density classification for the SRF study area.....	64
19.10	Figure 10. Gross vegetation classification for the SRF study area	65
19.11	Figure 11. Dominant ground surface features for the SRF study area.....	66
19.12	Figure 12. Hierarchical Landscape Classification of the SRF study area.....	67
19.13	Figure 13. Statistical Cluster Classification of the SRF study site.....	68
19.14	Figure 14. Example dry month and wet month water condition maps	69
19.15	Figure 15. Index of water level condition, SRF study area	73
19.16	Figure 16: Plots of wading bird observations for four time points representing four very different spatial patterns.....	75
19.17	Figure 17: Time series plot of all species estimated abundance and 95% confidence bars	75
19.18	Figure 18: Time series plot of long legged tactile feeders estimated abundance and 95% confidence bars. Time series plots of individual species in this guild are also provided.....	75
19.19	Figure 19: Time series plot of long legged visual feeders estimated abundance and 95% confidence bars. Time series plots of individual species in this guild are also provided.....	75
19.20	Figure 20: Time series plot of short legged visual feeders estimated abundance and 95% confidence bars. Time series plots of individual species in this guild are also provided.....	75
19.21	Figure 21: Time series plot of small dark and small white heron estimated abundance and 95% confidence bars	75
19.22	Figure 22: Water condition fractions pattern as a simplex	75
19.23	Figure 23: Weighted kappa statistics on water condition associations.....	75
19.24	Figure 24: Estimated extradispersion by date for all species combined.....	75
19.25	Figure 25: Estimated extradispersion by date for long legged tactile feeders.....	75
19.26	Figure 26: Estimated extradispersion by date for long legged visual feeders	75
19.27	Figure 27: Estimated extradispersion by date for short legged visual feeders.....	75
19.28	Figure 28: Estimated extradispersion by date for small dark and small white herons...	75
19.29	Figure 29: Map of the Everglades area, showing the boundaries of protected areas.....	75
20	TABLES.....	76
20.1	Table 1: Estimated percent observed by species for SRF pilot study.....	76
20.2	Table 2: Land element themes defined from existing maps	77
20.3	Table 3: Land element categorizations based on aerial observations	78
20.4	Table 4: Landscape classification based on multivariate analysis of land element themes	79

20.5	Table 5: Landscape classification based on multivariate analysis of land element themes 80	
20.6	Table 6: Models used to examine various dependence structures with the bird count data	81
20.7	Table 7: Schwartz's Bayesian Criterion values for alternate GLMM models for selected dates, all species combined	82
20.8	Table 8: Total counts and percentages by feeding guild and species	83
20.9	Table 9: Minimum and maximum estimated species abundance by feeding guild and species.....	84
20.10	Table 10: Jaccard index for association, averaged over time, between guilds, with associated variance of this index over time measured as the standard deviation.....	85
20.11	Table 11: Jaccard index for association, averaged over time, between species, with associated variance of this index over time measured as the standard deviation.....	86
20.12	Table 12: Decomposition of total variability in bird counts.....	88
20.13	Table 13: Importance of land element themes in describing spatial pattern	89
20.14	Table 14: Patterns of significant over- and under-utilization of cells with different land element themes	90
20.15	Table 15: Importance of HLC and SCC in describing spatial patterns.....	91
20.16	Table 16: Importance of surface water condition in describing temporal and spatial patterns.....	92
20.17	Table 17: Patterns of significant over- and under-utilization of cells with surface water classifications	93

2 Executive Summary

Foraging patterns of colonial wading birds (*Ciconiiformes*) in the 4300 sq km of the southern Everglades in Florida are described for 93 sampling dates during 1985 to 1995 based on observations from a systematic aerial sampling program. Sampling occurred primarily during the winter months (Dec-May) for each year with a mid-summer (August) sample. Additional months were included at the beginning of the program as well as just after hurricane Andrew. An index of bird abundance is used to examine temporal and spatial patterns as well as explore correlations with landscape and surface water coverage patterns. Location and counts of Great Egret (*Casmerobous albas*), Snowy Egret (*Egretta thula*), Great Blue Heron (*Ardea herodias*), Great White Heron (*Ardea herodias occidentalis*), White Ibis (*Eudimicus alba*), Glossy Ibis (*Eudimicus niger*), Wood Stork (*Myrcia americana*) and Roseate Spoonbill (*Ajaj ajaj*) were collected. Feeding guilds were created by pooling the above into the following: short-legged, tactile feeders (White Ibis, Glossy Ibis), long-legged, tactile feeders (Wood Storks, Roseate Spoonbills) and long-legged, visual feeders (Great Blue Herons, Great White Herons, Great Egrets).

The estimated numbers of wading birds demonstrated a peak in January, 1993 with 90,035 birds, with a minimum estimated number of 1327 birds for May, 1990. All species and feeding guilds experienced declines from December to May, although the year-to-year pattern was not always consistent. Only a fraction [actually a large fraction] of this temporal pattern can be attributed to hydrology changes, although it is clear that bird abundance was positively correlated with the fraction of total area classified as being of intermediate between wet or dry. Spatial patterns of bird occurrence was highly correlated with bird abundance which was also highly correlated with water condition changes which were themselves correlated to vegetation and other landscape patterns [so no need to include them all...]. [Change a lot here!]

3 Preface

The problem of wading birds in the Everglades is not a simple one. To understand why there appear to be so many fewer birds now than in historical times, we will have to consider many aspects of wading bird biology, as well as recent alterations to the Everglades environment. This report, with its associated publications, represents only the first step in that understanding. It deals with wading bird distribution and abundance, as monitored by the Systematic Reconnaissance Flights. It does *not* deal directly with wading bird breeding success. The reason is simple. We are uncomfortable relying on published summaries. But the raw data on wading bird breeding, while extensive, are highly fragmented. To locate and compile them is a large project unto itself. As this report is being produced, just such a project is underway. The results of analysing the compiled data on wading bird breeding success will provide another piece of the puzzle.

Note, though, that wading birds nest during the dry season of the Everglades (December–May). Their distribution over the landscape during this period represents the intensive foraging required to nest successfully in a limited window of time. Failure to forage effectively is cited a principle cause of nesting failure (Kushlan *et al.* 1975; Ogden *et al.* 1980; Frederick and Collopy 1988). So our analyses here are indirectly linked to breeding success.

This report is the result of many different inputs over a considerable period of time. Much of it is compiled from various unpublished drafts. Analyses from these drafts are described, even though not all of them led to usable results. The most recent work is Chapter 11, which describes a predictive model of wading bird abundance and distribution, and which will provide the

DRAFT COPY

foundation of a manuscript submitted for publication.

3.1 *A note on the ordering of Figures and Tables*

Many of the figures and tables in this report were prepared, with their accompanying legends, a considerable time before the final version of the text. Subsequent reordering of the layout of the report caused the figures and tables to be cited in a different order from that originally envisaged. To avoid unnecessary and time-consuming recreation, we have used the original material and retained the original reference information. Thus the figures and tables are at the end of the report with their original numbers, and in their original order (Figure 1, then Figure 2, and so on) for easy location. Citation in the text, while not in the original order, still uses the original numbers.

4 Introduction to the Everglades

4.1 *Location of the Everglades*

For the purposes of this report, we consider the Florida Everglades to be the area of Florida south of Lake Okechobee that is currently, or was historically, wetland. Currently, there are three major regions comprising the Everglades, as detailed in Figure 29.

- 1) *The Water Conservation Areas (WCAs)*. These occupy the area immediately south of Lake Okechobee, and extend south to the Tamiami Trail.
- 2) *Big Cypress National Preserve (BCNP)*. This is an area to the west of the WCAs. It extends south to just below the Tamiami trail.
- 3) *Everglades National Park (ENP)*. This is the area south of the other two. It extends from Tamiami Trail in the north to the coastal mangroves in the south, and from the edge of the Maimi urban area in the west to the coastal mangroves in the east.

This report is concerned primarily with Everglades National Park. The data available to us at the time of the report come from that region alone. We recognise that the Everglades as a system encompasses all the areas above, and will address this issue at the close of Chapter 11, as well as elsewhere.

4.2 *Location of the Study Site*

The study area consisted of 4300 km² of the southern Everglades, including all of Everglades National Park and the southern part of Big Cypress National Preserve, but excluding Florida Bay and the immediate west and southern coastal edge (Figure 2).

Books have been written about the biotic and abiotic composition of the Everglades so no attempt will be made here to describe the area in detail. In general, the southern Everglades is located in the southeastern coastal plain of the North American continent at the southern end of the Florida peninsula. It is the southern-most part of the larger, Kissimmee-Okeechobee-Everglades regional system. Geologically, this system is very young, being composted of peat and marl deposited on limestone bedrock over the past 5000 years (Brown 1988). This bedrock is depressed in the center of the system, limiting potential freshwater drainage to the Atlantic Ocean and producing marshes, called sloughs, through which freshwater flows and empties into Florida Bay. Topography, water depth and water flow combine to produce water lily sloughs and tree islands in the center portions of the system, with sawgrass and mixed sawgrass/wax myrtle in the higher areas. The coastal fringe vegetation is dominated by mangrove.

The climate of the southern Everglades is humid subtropical, being affected strongly by the Gulf Stream and wind patterns of the Atlantic Ocean and Gulf of Mexico. It experiences mild temperatures year-round with distinctive wet and dry seasons. Summer is wet with almost daily convective rainfall. Winter is dry, with what little rainfall occurs being the result of continental frontal systems reaching down into the tip of Florida. Drought, fire, periodic freezes, and tropical storms cause major short-term changes in abiotic and biotic components of the system (Duever *et al.*, 1994, Gunderson and Snyder, 1994). Human changes in the regional hydrology (Light and Dineen, 1994), agriculture (Snyder and Davidson, 1994), and climate induced sea level rise (Wanless *et al.*, 1994) have changed much of the Everglades system from what it was just 100 years ago.

4.3 Climate

Although the northern region of the Florida peninsula is classified as temperate continental, the southern end that comprises the Everglades regional landscape is subtropical. A broad transitional zone occurs between the two distinct regions. The southern, subtropical region is located at the convergence of the Atlantic Ocean and the Gulf of Mexico. These large bodies of water, their defining currents (e.g., the Gulf Stream's proximity to the southern tip of the Florida peninsula), and the periodic occurrence of El Niño years affect the region's climatic regimes.

4.4 Physiography and geology

The Everglades region is located at the southern tip of the low-lying Florida peninsula, an exposed portion of the broad Florida Shelf, separating the Atlantic Ocean from the Gulf of Mexico. The entire peninsula lies within the Southeast Coastal Plain of North America. Global climate change has raised sea level above the elevation of the peninsula several times since its original formation over 25 million years ago. Submergence of the southern portion of the peninsula has occurred with each transgression of the sea, followed by reemergence of the peninsula during periods of global cooling, glacial advance, and subsequent sea level fall. The latest emergence of the peninsula above sea level occurred approximately four to five thousand years ago (Brown 1988).

At the present level of the (rising) sea, approximately XX percent of the contemporary or post-drainage Everglades regional landscape is upland, XX percent freshwater wetland and approximately XX percent estuarine (i.e., coastal estuaries, bays, shallow lagoons and reef tracts). The underlying geology of the contemporary landscape reflects a history of recurrent submergence and emergence. Underlying limestone bedrock is predominately marine in origin (i.e., an ancient coral reef). Variation in the types, areal extent, and depth of overlying substrates that occur throughout the regional landscape reflect geologic and climatic history (Brown 1988), as well as recent past water management practices and related land-use patterns in the region.

Prominent physiographic features of the upper reaches of the regional floodplain include the Kissimmee River basin that drains into the Lake Okeechobee basin from the north; the Catalahotchee River basin, draining to the west of the lake into the Gulf of Mexico; and the St. Lucie basin that drains to the east into the Atlantic Ocean. These drainages (and related canal systems) transport freshwater, interior flows from upper reaches of the floodplain to tidewater. The flat floodplain (wetland) landscape of the Everglades lies immediately downstream to the south, flanked by high elevation, ephemeral or dry prairies to the east and west, a forested coastal ridge to the east along the Atlantic Ocean, a low lying ridge adjacent and to the west of the southern, interior wetlands and a forested swamp further to the west. The Atlantic, Florida Bay and Gulf coastlines comprise the southeastern, southern and southwestern, respectively, border of the region's mainland area.

DRAFT COPY

This wetland landscape is further dissected by four major watersheds that transverse the major physiographic sub-regions (forested swamp, grassland savannah and forested estuaries) comprising the southern Everglades. The most western watershed is the Big Cypress drainage that drains water flow from the forested swamp to the Gulf coast. Shark River Slough, a central flowway in the downstream reaches of the floodplain, is the major watershed of the southern Everglades that drains water flow from the interior wetlands into the mangrove estuaries of the southwestern Gulf coast. A smaller flowway (Taylor Slough), adjacent and to the east of this central flowway, drains water flow from the eastern transitional, ephemeral wetlands into the estuaries along the Florida Bay coastline. East of the Taylor Slough drainage, and extending to the Atlantic coast, is the Eastern Panhandle drainage system that drains water flow from the Atlantic coastal ridge to upper reaches of the Taylor Slough drainage that occur to the west and to forested mangrove estuaries along the Atlantic coast to the east and southeast of this ridge or divide. An ancient beach rim separates the coastal estuaries of the region from the freshwater prairies of the interior and adjacent forested swamp to the west of these prairies.

Elevational gradients in the freshwater wetlands are slight, averaging approximately 3 cm per km from the upper to the lower reaches of the floodplain (Parker 1974). Elevation ranges from approximately 6 m msl in the northern reaches of the floodplain (south of Lake Okeechobee) to sea level along the Gulf of Mexico and Florida Bay coasts. Elevational gradients are relatively steep in the northern versus southern reaches, creating relatively low, hydrologic gradients and the natural pooling of surface water, in the extreme downstream, interior portions of the floodplain. The central flowway or slough in the interior of the southern portions of the floodplain is flanked by the high elevation, severely drained, limestone escarpments of the eastern transitional wetlands and by transitional wetlands of intermediate hydroperiods to the west.

Before severe drainage, numerous depressions or solution holes in the exposed, highly eroded bedrock of the eastern, high elevation marshes retained ground water throughout most dry seasons, providing dry season refugia for aquatic organisms throughout the eastern transitional wetlands, as seasonal surface water receded across the landscape. Most solution holes in the highly eroded bedrock underlying the high elevation, transitional wetlands to the west, however, are overlain by sands and sandy marls. Small, parallel ridges in the topography extend throughout these short-hydroperiod wetlands, creating intermittent, longitudinal pools of surface water. During most dry seasons, however, these pools are ephemeral and quickly disappear as water levels progressively recede across the landscape.

Open water ponds or "alligator holes", formed by excavating activities of alligators in marsh areas characterized by deep muck or sand deposits, occur throughout most of the interior, freshwater wetland. However, alligator occupancy in the drained, upstream reaches and in downstream reaches of the impounded marshes of the interior (where deep, semi-permanent pools of surface water have formed) is low. Such extreme hydroperiod ranges provide unsuitable habitat for year-round occupancy by alligators (Fleming in prep). Consequently the occurrence of such "alligator holes" in the upper and lower reaches of these impounded marshes of the northern Everglades is rare.

Elevational gradients of the mixed cypress/pine forested swamp, adjacent and to the west of the interior floodplain, are also relatively steep. They range from approximately 6 to 18 cm per km from the swamp interior to headwater regions of downstream estuaries (Carter *et al.* 1973).

General elevation ranges in downstream estuaries are slight (less than 0.3 m) from headwater or upper to lower estuarine zones. Along the Gulf coast, numerous tidal creeks and tributaries in the upper estuarine zone converge to form tidal rivers, or channels, with steep banks and strong inter-tidal currents. Widespread, shallow tidal flats border these tidal rivers or channels. Such

DRAFT COPY

well-developed tidal channels or rivers are absent in the estuaries of Florida Bay, where numerous small tidal creeks and tributaries characterize the drainage patterns. In addition, deep water tidal bays and inland lakes occur in the middle and lower estuarine zones along both the Gulf and Florida Bay coastlines.

4.5 Soils

Limestone bedrock (an ancient coral reef) underlying the wetland floodplain is overlaid with deep muck deposits in the northern reaches, graduating to shallow deposits in the southern, interior reaches. Exposed, highly eroded, limestone bedrock of high elevation marshes that border these southern interior wetlands is partially overlaid by organic and inorganic marls to the east and sandy marls to the west. Fine sands graduating to sandy marls and localized muck deposits underlie the cypress swamp to the west. Mixed, unclassified tidal marls and mucks underlie the coastal estuaries to the south. These substrates extend seaward of the ancient beach rim and were formed by the deposition of sediments from upstream reaches that were transported by water flow over this ancient beach rim (during the last episode of global cooling, glacial advance, and falling sea level). These coastal, sediment banks or "mud flats" are now colonized by mangrove plants or forests. Submergence and/or sea level rise, however, have again started to erode the coastal berms and interior portions (lower and middle reaches) of these mangrove-colonized, coastal flats.

Human alterations in the distribution, volume and timing of natural patterns of water flow across the landscape have also changed soil formation processes. In areas of severe drainage, oxidation and/or the burning off of underlying organic (muck) soils by intense wildfires, has resulted in the subsequent formation of more inorganic (marl) soils or the complete loss of soil (refs).

4.6 Vegetation

Plant communities of the region are comprised of tropical species. Dominant vegetative cover types in the freshwater wetlands include extensive, dense sawgrass prairies interspersed with tree islands and wet prairies in the northern reaches of the study area. A much finer-scale mosaic of open and dense sawgrass, tree island prairies, wet prairies, open water areas and sloughs characterizes the lower hydrologic gradients of the interior, freshwater wetlands of the southern portion of the floodplain. These interior wetlands are bordered by remanent, high elevation, dry prairies to the east and west. Since drainage of these high elevation wetlands, several exotic plant species (*Melaluca*, *Casarina* [sp?], and *Shinus* spp.) have become established and continue to expand their geographic range across the wetland terrain despite on-going control or eradication efforts.

The mixed-cypress/pine forested swamp, adjacent and to the west of the interior wetlands, is typically comprised of pine forests on high elevation sites; cypress strands dominate in shallow sloughs and depressions in the ground surface; and mixed cypress-pine forests are the dominate cover type in areas of intermediate, relative elevations. Interspersed throughout the higher elevations are dry prairies.

Dominate cover types of the downstream coastal estuaries include tall, fringing mangrove forests along the banks of tidal rivers and the shoreline of the Gulf coast. These forests are interspersed with isolated, coastal prairies that are being colonized by the surrounding mangrove communities. [Both the fringing mangrove forests and the coastal prairies are also being invaded and colonized by the exotic Brazilian Pepper (*Shinus* spp.).] Mangrove forests of variable height and open grasslands or prairies also characterize the headwater or upper Gulf estuarine zone. Mixed, short to tall mangrove forests dominate the lower and middle estuarine zones of Florida Bay, graduating to short or dwarf mangrove forests in upper reaches. Mixed, short to tall

mangrove forests and extensive open mud flats characterize the convergence of the Florida Bay and Gulf coastal estuaries at the Cape Sable peninsula.

4.7 *Animals*

A diverse mixture of sub-tropical and temperate animal communities inhabit the wetland floodplain. These communities comprise terrestrial-based (macrophytes/seeds) and more aquatic-based (detritus/periphyton) food webs in the upper versus lower, respectively, elevational gradients of the floodplain.

The various, dominant plant communities of the regional floodplain do not provide high quality "green" forage or primary production. However, such dominants do provide an abundant source of detrital or "brown" organic material that serves as a food base for the aquatic food webs. Consequently, more wetland-dependent or aquatic animal life comprise this "brown" or aquatic food web (i.e., aquatic invertebrates (snails, crawfish and prawns), fishes, amphibians (frogs and salamanders), small reptiles (aquatic turtles and snakes), crocodilians (crocodiles and alligators), a diverse array of wading birds, Snail kites, Anhingas, and otters). These species or functional groups of species dominate the trophic structure of the native fauna in the (wetter) relatively lower, elevational gradients of the freshwater, interior and coastal estuarine portions of the seasonally inundated floodplain.

A more terrestrial-based food web (i.e., terrestrial insects, crawfish, lizards, turtles, snakes, mice, rats, muskrats, marsh rabbits, deer, mink, raccoons, bobcat, panther and bear) characterizes the (drier), relatively higher elevations or gradients of the region—the tree island prairies of the interior, the drier prairies of the remaining ephemeral wetlands and upland forested regions of the eastern coastal ridge and pinelands, as well as the mixed cypress/pine forested swamp land west of the interior prairies.

4.8 *Hydrological history of the Everglades Landscape*

4.8.1 *Pre-drainage*

The historical or pre-drainage Everglades regional landscape (circa mid-1800s) was expansive in areal extent, encompassing approximately 47 500 km². A silt-laden, feeder river (Kissimmee River) drained into a large inland lake (Lake Okeechobee) located at the headwaters of a downstream, flat floodplain (wetland) landscape. Overflow from the lake during the rainy season, and groundwater seepage year round, contributed water flow to adjacent, downstream reaches of the floodplain. Outflows at the southern end of the floodplain were only 15% of inflows, due to a high evapotranspiration rate, and terminated in a salt environment (i.e., into the sea). Although the feeder river to the floodplain was silt-laden, the system was oligotrophic due to blockage of flow at the junction of the feeder river and floodplain by the lake berm, causing a decrease of flow rates, and the deposition of suspended sediments/nutrients before flows reached the downstream floodplain.

The landscape throughout was relatively flat (very little topographic relief and low hydrologic gradients). Slight changes in elevation were characterized by different periods of inundation (hydroperiod) and were therefore dominated by different soil types and plant communities. This highly variable microtopography created a fine-scale mosaic of wetland types. Accordingly, the landscape was characterized by complex patterns of spatial heterogeneity (and temporal variability).

A forested ridge occurred to the east of the floodplain, separating the freshwater wetlands from the Atlantic Ocean. Extensive fire-adapted, open pineland, interspersed with tropical hardwood hammocks, characterized higher elevations of the ridge. Transverse glades or lower elevation

flowways, underlain by mixed marl/muck soils, penetrated through the ridge and drained to the coast.

An extensive, transitional area of high elevation, ephemeral (short-hydroperiod) wetlands, characterized by rock and marl soils, separated these eastern pinelands from the interior (long-hydroperiod) wetlands and central sloughs of the region. High elevation marshes also occurred to the west, separating the interior wetlands from a mixed cypress/pine forested swamp that extended further west to the edge of the floodplain along the Gulf coast. Predominate soils of these western, high elevation wetlands and adjacent forested swamp consisted of course to fine sands, sandy marls and marls.

The interior wetlands and central sloughs were underlain by muck soils. Numerous tree islands—tropical hardwood hammocks, swamp forests, willowheads, bay heads, cypress domes, etc.—occurred throughout a variety of wetland types comprising these interior marshes.

Mangrove forests dominated the coastal estuaries that occurred downstream to the southeast, south and southwest of these freshwater wetlands. Originally formed by sediments transported from upstream freshwater to downstream estuarine reaches, substrates of these estuaries were primarily tidal marls and mucks.

The natural hydrologic regime of the region was characterized by pulse (seasonal) flows (wet and dry season) and uninterrupted wet season sheet flow across the floodplain during the apex of flows. Such flows inundated the extensive areas of ephemeral wetlands for 3–6 months of most years. Attenuated seasonal changes in marsh water depths, caused by slow flows due to the low hydrologic gradients of the landscape, resulted in persistent dry season flow from upstream to downstream reaches of the interior wetlands and coastal estuaries (Perkins *et al.* 1994).

The seasonal recession of water levels across the landscape created a migratory drying front or transitional ecotone or edge that could expose the aquatic prey to predation by top-level carnivores (Harris *et al.* 1996). This dynamic expansion and storage of sheet flow during the wet season, followed by lagged flows from upstream to downstream reaches throughout most dry season months, permitted long intervals (9–12 years) between the occurrence of major drydowns in the interior wetlands and central sloughs, and large, dry season freshwater flows into the estuaries (Perkins *et al.* 1994).

The landscape was characterized by top level, long-lived, large, kilothermic predators (crocodilians) and a diverse herpa- and avi-fauna. It was also an important point of sequestration for either migratory or breeding species, particularly for wading birds (refs).

The hydrologic regime (distribution, volume and timing of water flow) was a major driving force of trophic interactions and trophic structure. Annual variation in rainfall and related marsh water depths provided alternating years more favorable to aquatic organisms when rainfall/runoff was above normal, and more favorable years for semi-aquatic and terrestrial organisms during years of less rainfall. The wet season months of peak flows were normally a time of dispersal, recolonization of seasonally inundated marsh, and peak production of aquatic organisms. Dry season months were normally a time of consumerism of these aquatic organisms by top-level predators (e.g., wading birds, crocodilians, etc.).

Episodic events (periodic freezes during the winter dry season, lightning-ignited wildfires during the late dry and early wet season months, and hurricanes during the summer wet season) furthered sculpted the landscape. Such events were important perturbations at the micro-habitat scale or individual level, and important landscape processes and regional phenomena at the macro-habitat scale or population level, driving adaptation and evolution of the characteristic biota.

4.8.2 *Post-drainage: General*

The Florida Everglades is one of the world's largest and best known wetlands. Arguably, the impacts of humans on such a globally significant wetland in the immediate past century rival no other place on earth in altering the fundamental surface structure, and fundamental energy and material processes faster or more dramatically than that of south Florida. The staggering rates of human population growth, wetland drainage and deforestation, for example, make south Florida a global-level experiment that challenges present legal structure and jurisprudence, technological capacity, and even environmental science conceptual frameworks. The rapidity and magnitude of change, and loss of biodiversity that is presently occurring in south Florida, are daunting challenges to maintaining the integrity of this globally significant wetland. Defining solutions for maintaining the integrity of the greater Everglades system is a high priority for many federal, state and local governmental agencies as well as conservation groups.

The construction of canals, levees and roadways across the entire south Florida terrain have created profound changes in the seasonal and annual patterns of wetting and drying of the landscape. Drainage, and the subsequent expansion of agriculture from high elevation, coastal ridges to the wetland interior, as regional water levels are continuously lowered over time, and the "inevitable" progressive pattern of development of these agricultural lands into residential and then urban areas, also continues at an accelerating rate. As a result of this land transformation process, one of the world's most significant wetlands is now juxtaposed next to one of the western hemisphere's most vibrant and fastest growing cities.

The natural water flow within the Everglades region was altered by the construction of major canals draining Lake Okeechobee and northern parts of the Everglades in the early nineteen hundreds. These canals, completed and fully operational in 1921, affected the hydrology of the Everglades regional landscape, including both hydroperiods and water levels (Leach *et al.* 1972).

Subsequent human settlement of the region has also brought with it a fifty percent reduction in the areal extent of the wetland system and a related loss of heterogeneity or diversity of wetland types. This extensive habitat loss occurred disproportionately, in upland habitats and ephemeral wetlands, particularly along the entire eastern edges of the floodplain. As a result of this extensive habitat loss, or severe truncation of the historical floodplain, protected natural wetland areas of the post-drainage landscape now include only the lower elevations of the historical and expansive floodplain that once encompassed a wide range of environmental gradients (Fleming *et al.* 1995).

Continual, widespread drainage, landscape fragmentation by the construction of canals and roadways across the wetland terrain, isolation of remaining natural wetlands, deforestation, alteration of the natural hydrologic regime by past and on-going water management practices, related changes in salinity and fire regimes, eutrophication, contaminants and the invasion of non-indigenous species have all resulted in faunal collapse and erosion of native biodiversity in remaining protected natural areas.

Such profound changes are altering both the contextual surroundings and content of the historical Everglades landscape, resulting in landscape fragmentation and isolation, related faunal collapse and the erosion of biodiversity in remaining protected natural areas (Harris *et al.* 1994). The effects of this changing Everglades landscape on the distribution, abundance and diversity of the regional wading bird population are particularly foretelling.

Historically, changes in the natural hydrology of the Everglades system have produced severe declines in all feeding guilds of migratory, or overwintering, colonial wading bird populations in south Florida (Robertson and Kushlan 1974; Kushlan and White 1977; Kushlan *et al.* 1984;

DRAFT COPY

Bancroft 1989; Powell 1987; **Powell *et al.* 1989**). These more recent declines follow a recovery in wading bird numbers from a turn of the century decline that was due to unregulated hunting of the wading birds for their plumes. Severe reductions (90% or more) in the number of nesting colonial wading birds have also been reported in the last two decades (Robertson and Kushlan 1974; Kushlan and White 1977; Ogden 1978; Kushlan *et al.* 1984; Frederick and Collopy 1988; **Powell *et al.* 1989**). The mechanisms relating declines in Everglades wading bird populations to the changes brought about by canal construction, drainage, and changes in regional water management are the subject of continuing debate.

The historical distribution, volume, timing, areal extent, continuity and periodicity of water flow across the landscape have been greatly altered throughout the entire floodplain by the construction of canals and levees, and the subsequent introduction of water management regulation.

4.8.3 *Post-drainage: Freshwater wetland hydrodynamics*

Water flow is no longer driven by the naturally occurring, seasonally-pulsed flows that characterized the historical, hydrologic regime. Instead of attenuated, seasonal changes in marsh water depth, most water flow from upstream to downstream reaches is through canal water deliveries, with greatly increased flow rates regulated by numerous water control structures. Impoundment of flow, as well as increased and prolonged rates of drainage, have created areas of hydroperiod extremes from the peripheral edges of the floodplain to the interior wetlands and central sloughs. Such blockage or fragmentation in the continuity of water flow has also interrupted the continuity of water flow among major drainage basins, watersheds, physiographic sub-regions and terrain patterns of the southern Everglades.

4.8.4 *Post-drainage: Coastal (estuarine) wetland hydrodynamics*

Freshwater flows from upstream reaches of the floodplain drain into the coastal estuaries through numerous, small creeks and tributaries originating in the ecotonal, headwater region. This region is delineated by the ancient beach rim that separates the upstream, freshwater wetlands from the coastal estuaries. A brackish water zone or front characterizes this ecotonal region. This brackish front migrates from upper to lower reaches of the estuaries in response to local rainfall/runoff and the seasonal pulse of freshwater flows that drain into the estuaries from upstream, freshwater reaches of the floodplain. The creeks and tributaries of this headwater region join further downstream to form tidal rivers with steep banks and strong, inter-tidal currents that drain into the Gulf through passages in coastal beach rims. Tidal mud flats that occur from upstream to downstream reaches between these tidal rivers are normally inundated by local rainfall/runoff during the wet season months and, in middle to upper reaches of the estuary, by water flow over the banks of these rivers at high tide only during the dry season. An annual decrease of sea level of approximately 1 foot msl in the Gulf during the winter, dry season, however, allows drainage of the mud flats to occur during the dry season, except for periods of high, tidal overflow. In addition, strong, onshore winds (associated with the passing of winter cold fronts from the north) along the Gulf and Florida Bay coastlines accentuate high, tidal overflow of these river banks and subsequent inundation of adjacent, tidal flats. Decreased, dry season freshwater flows under post-drainage conditions have also increased salinity ranges throughout these estuaries.

4.8.5 *Overview*

Changes in the system's seasonal hydrology have formed semi-to-permanent deep water pools in impoundments constructed to the north of this downstream area. This impoundment or fragmentation of the continuity of water flows has also altered marsh hydroperiods and drying patterns downstream of these impoundments throughout the entire southern Everglades (Leach

et al. 1972). It has magnified the reversals in the drying pattern in the southern Everglades that occur in response to management discharges of excess water from upstream reaches of the floodplain. There has also been a westward shift in water flow outside of the historical flowway, Shark River Slough, of the southern Everglades (Johnson *et al.* 1993). A subsequent decrease in dry season flows through this slough (Johnson *et al.* 1993) is credited with an increased frequency, magnitude and duration of major drydowns in the slough and a concurrent reduction in the areal extent of surface water pools in the headwater region of Gulf coast estuaries immediately downstream (Craighead 1968). Finally, severe drainage of high elevation marshes that flank this central flowway of the southern Everglades, particularly along the eastern edge of the floodplain, has been documented (Johnson and Fennema 1989; Johnson *et al.* 1993; Van Lent 1993).

4.9 Related changes in fundamental landscape characteristics

Human settlement of the region has altered a number of fundamental landscape characteristics. Such alterations have occurred as a result of:

- 1) widespread deforestation along the northern and entire eastern perimeter of the floodplain (refs), following
- 2) drainage, and subsequent settlement of the Atlantic coastal ridge and the adjacent, high elevation marshes; such drainage has resulted in a disproportionate reduction in shallow water areas of the floodplain (the ephemeral wetlands) that were inundated during the late wet/early dry season (a time of year when regional water levels normally are highest) (Fleming *et al.* 1994);
- 3) alteration in the periodicity of water flow throughout the southern Everglades floodplain (Perkins *et al.* 1994), and a resultant increased frequency, magnitude and duration of major dry downs in the interior wetlands and sloughs (Perkins *et al.* 1994), creating adverse interannual effects on aquatic production (Loftus and Eklund 1994);
- 4) related changes in the fire and salinity regimes since the introduction of water management regulation (refs);
- 5) the invasion of non-indigenous species (refs);
- 6) human enrichment of formally oligotrophic waters (refs), and
- 7) the introduction of contaminants and mercury in the food webs of the remaining protected natural areas, as a result of airborne and upstream, runoff deposition of nutrients and contaminants from adjacent, settled (agricultural, residential, urban and commercial) portions of the historical floodplain (refs). High levels of mercury now occur in wading birds and other top-level carnivores (refs).

4.10 Animal issues of concern

Animal issues of concern include:

- 1) the overall decline in historical numbers of native and endemic taxa, particularly the chronic decline in colonial wading birds across all feeding guilds;
- 2) changes in the historical patterns of the spatial and temporal distribution of species;
- 3) resultant changes in patterns of community structure; and
- 4) related changes in the pattern and magnitude of energy and material flows.

Most animal populations in the region are now characterized by reduced densities and biomass,

increased mortality, reduced reproductive potential and juveniles survivorship, and changes in persistence and recoverability from disturbances (Fleming *et al.* 1997). These characteristics are usually associated with fragmented, natural landscapes in which populations are frequently affected by environmental stochasticity, demographic stochasticity, social dysfunction or genetic deterioration (Noss 1987). Many characteristic Everglades species are now rare, threatened or endangered, with a concurrent increase in weedy or common species (generalists) and invasion of non-indigenous or exotic species (Harris *et al.* 1994).

4.11 *Historic and present status of colonial wading bird populations*

Current consensus has it that the historical or pre-drainage, Everglades landscape had three features considered crucial to wading bird survival and successful recruitment. First, during the wet and early dry season, shallow water inundated extensive, ephemeral wetlands. Second, dry-downs of the interior wetlands and central sloughs during the dry season were infrequent, since delayed flows from upstream to downstream reaches buffered seasonal water depth changes in these interior marshes. Third, large freshwater flows through the entire system into the coastal estuaries occurred even during dry season months (Parker 1974; Smith *et al.* 1989; NPS 1990; Perkins *et al.* 1993; Johnson *et al.* 1993).

Wading birds that forage and attempt to nest at traditional colony sites in the southern Everglades have demonstrated poor reproductive performance under post-drainage conditions (Kushlan and White 1977; Frederick and Collopy 1988; Powell 1987; Bancroft *et al.* 1991). It has been hypothesized that wading birds require progressively drying conditions, i.e., a receding drying edge, during the dry season to concentrate prey to sufficiently high densities for efficient foraging. Documented changes in the pattern of drying, as indexed by drying rates or changes in water levels measured at hydrological stations, have been correlated to nesting delays of Everglades wading birds with subsequent poor reproductive success and fledgling survival (Kushlan *et al.* 1975; Ogden *et al.* 1980; Frederick and Collopy 1988). Comparisons of drying rates between years, however, has not provided a useful index of marsh water depth changes, probably because water levels at the beginning of the dry season differ so much from year-to-year. Comparisons among years has also failed to take into account changes in upstream flow distributions which can affect the drying rates at downstream hydrological stations. These analyses using drying rate have also ignored the spatial variability of drying patterns in a landscape.

Concurrent with alteration of the natural hydrology of the system, severe declines occurred in migratory, or overwintering, colonial wading bird populations in south Florida, across all feeding guilds (Robertson and Kushlan 1974; Kushlan and White 1977; Kushlan *et al.* 1984; Bancroft 1989; Powell 1987; Powell *et al.* 1989). These concurrent declines followed a recovery in wading bird numbers from an initial decline that was due to unregulated, marketing hunting of the wading birds for their plumes during the turn of the century. Severe declines (90% reduction) in the number of nesting colonial wading birds have also occurred concurrent with alteration of the natural hydrologic regime (Robertson and Kushlan 1974; Kushlan and White 1977; Ogden 1978; Kushlan *et al.* 1984; Frederick and Collopy 1988; Powell *et al.* 1989). The specific mechanisms relating declines in Everglades wading bird populations to the changes brought about by canal construction and water management in the region is the subject of continuing debate.

Landscape characteristics of the post-drainage Everglades are not meeting the habitat requirements of wading birds to maintain viable breeding populations. Wading birds that forage and attempt to nest at traditional colony sites in the southern Everglades have demonstrated poor reproductive performance under post-drainage conditions (Kushlan and White 1977; Frederick and Collopy 1988; Powell 1987; Bancroft *et al.* 1991).

Numerous hypotheses have been advanced and debated concerning the causal mechanisms associated with alteration of the natural hydrologic regime of the regional floodplain that may explain the severe and well documented declines in both breeding and non-breeding wading birds in south Florida (refs). Hypotheses regarding such mechanisms that are either directly or indirectly relevant to the data collected and analyzed during this study, particularly for the downstream reaches of the floodplain that comprise the southern Everglades, include:

1. Wading bird declines are directly related to loss of overall foraging habitat;
2. Wading bird declines are directly related to the disproportionate loss of specific (seasonal), foraging habitat types;
3. Wading bird declines are directly related to changes in the spatial configuration of foraging habitat types;
4. Wading bird declines are indirectly related to fragmentation in the continuity of water flow among foraging habitat types;
5. Wading bird declines are directly related to increased, seasonal, water level fluctuations within foraging habitat types; and
6. Wading bird declines are indirectly related to alteration in the hydrologic periodicity (drydown frequency, magnitude and duration) of foraging habitat.

All the above hypotheses relate to changes in wading bird prey resources, seasonal prey availability, foraging patterns and bioenergetics of adults, timing of nesting and the related reproductive success of nesting pairs at traditional colony sites in the estuaries. We argue that it is the availability of shallow water to foraging wading birds that offers the best explanation for the observed shifts in feeding concentrations of most species.

Feeding pressure reduces prey density (Kushlan 1976a and b, 1978), hence a continual supply of new shallow water areas is required to sustain prey availability for wintering wading birds during the dry season. An elevated high drying rate is necessary to produce a moving drying front, and hence a continual supply of new available prey. However, too high a drying rate results in widespread, discontinuous surface water in the interior freshwater wetlands and in the estuaries, and the loss of this water during subsequent months. Such loss of surface water disrupts foraging activities of wading birds and forces their dispersal to other areas.

4.12 Project objectives (management questions)

Data from twelve consecutive years (1985–1996) of systematic aerial sampling of colonial wading bird populations in the southern Everglades will be used to examine how changes in wading bird abundance and spatial foraging pattern can be explained by post-drainage landscape characteristics and changes in hydrologic patterns. In particular, the temporal and spatial patterns of foraging colonial wading birds will be described and attempts made to relate these patterns to various classifications of the spatial landscape as well as temporal changes in the spatial pattern of surface water. The discussion will address the issue of foraging levels and progressive drying conditions.

Two non-analytic goals are the estimation of total numbers of birds and population estimates by class of bird, and the estimation/identification of the spatial (map) distribution of the birds.

5 The SRF Survey

5.1 *The SRF approach*

The Systematic Reconnaissance Flight (SRF) protocol is a set of standard aerial transect counting techniques employed within a systematic sampling design (Norton-Griffiths 1978). The protocol was used to obtain spatially explicit data on the distribution and abundance of various wading bird species. Other characteristics, such as species richness, accumulated species richness, species diversity, species relative density and biomass distributions can be derived from these data.

5.2 *History of the SRF project*

The original objectives of the SRF surveys were 1) to provide resource managers and Everglades researchers with maps of wading bird foraging distributions and, from the counts made, 2) to develop indices of total wading bird abundance with breakdown values by species and selected feeding guilds. Aerial systematic transect sampling was chosen for efficiency and economy, and to avoid the inefficiencies and observer problems associated with finding and observing a randomly assigned plots in a relatively uniform landscape. By orienting transects across the major topographical contours of the study area, total counts for each transect were expected to be relatively homogeneous, resulting in more precise final estimates. Each transect was further subdivided into 2km segments into which observations were recorded. These data were used to construct spatial maps of foraging distributions. These same datasets form the foundation for the statistical analysis.

It is important to point out that the SRF survey was designed *a priori* with certain types of analyses in mind. [continue]

5.3 *The SRF protocol*

The Systematic Reconnaissance Flight (Wading Bird) Survey utilizes the aerial transect method of counting birds. The principle is that an aircraft flies in a straight line from one side of the census zone to the other at a fixed height above the ground while observers count all birds within a strip of ground defined by markers on the windows and struts of the plane.

Fixed-wing, light aircraft with two bird observers, an observer of surface water condition, and a pilot, followed predetermined parallel flight lines (transects) oriented east-to-west across dominant topographical contours of the study area and spaced systematically from north to south across the study region. A preliminary study in 1984 indicated that there was little difference in total estimates and standard errors between surveys with lines placed 1 km versus 2 km apart, whereas there was significant loss of information with a less dense (≥ 3 km apart) design. Thus a 2 km spacing was used with each transect divided into 2 km sections (cells) for analysis and mapping purposes. A Loran C unit, calibrated from a known ground reference point before each survey flight, was used to provide flight line repeatability and to define line segments for recording wading bird counts and other observations. The sampling grid is overlaid on the study area in Figure 1. A 2 x 2 km grid cell, based on the 2 km separated sampling transects and associated 2 km line segments, was used to map wading bird counts.

The final sampling protocol was based on the results of a pilot study conducted in 1984 which examined the effect of observation height (45.7, 70 and 91.4 m), strip width (75, 150 and 225 m) and plane speed (80, 95, and 110 knots) on the wading bird counting accuracy of observers. These are known sources of counting bias (Norton-Griffins, 1974), and the goal was to minimize their effects in the final sampling protocol. The experiment used full factorial treatments with replicate blocking on day and time-of-day. Transects were run east-to-west across the direction of water

flow in the slough, in an attempt to avoid biases due to repeated sampling within very homogeneous areas or obvious effects of sun position on observability. Each transect contained both wet and dry areas, with habitat restricted to the slough areas. This avoided the more difficult mangrove habitat to the west and south of the study area.

Observer visibility bias was estimated for different size and color classes of wading birds using a modification of the stepwise, multiple-regression approach suggested by Caughley *et al.* (1976), as well as standard analysis-of-variance techniques. In general, there were no observer effects, as expected, since only experienced observers were used. Strip width, as a stand-alone factor as well as a component of an interaction with one or more of the other factors, was the major significant effect, followed by plane height. As with other applications of the SRF methodology, total model R^2 was very small indicating that there was a lot of response variability not explained by the factors and their interactions. The standard error of the predicted values was about 20%. For the sampling protocol used in the main study: 45m height, 80 knot ground speed, and 150m strip width (see Figure 2), the pilot experiment indicated that about 85% of total wading birds would be observed, with only about 30% of small dark birds actually present being counted (Table 1) [Note: Table 1 is blank. Do the numbers exist?].

Surveys were undertaken during the first week of each month, between 0800 and 1630 hrs at an approximate height of 45m, 80 knot ground speed, and with left and right observation strip widths of 150m resulting in a 15% systematic sample of the study area. Weather during the dry season, December to May, allowed three days for completion of each survey. This time period for observation was often extended during the wetter months due to convective storm activity. Flights were conducted on a monthly basis each year during the dry season, with one added mid-wet season flight (August). Flights were concentrated during the dry season because wading birds concentrate their reproductive activity during this period. In the first year of the study, flights were made monthly for the whole year, and additional flights were made in 1992 and 1993 following hurricane Andrew.

Location and cell count data of foraging colonial wading birds during each multi-species aerial survey were collected on the following species: Great Egret (*Casmerobius albas*), Snowy Egret (*Egretta thula*), Great Blue Heron (*Ardea herodias*), Great White Heron (*Ardea herodias occidentalis*), White Ibis (*Eudimicus alba*), Glossy Ibis (*Eudimicus niger*), Wood Stork (*Myrcia americana*) and Roseate Spoonbill (*Ajaj ajaj*). Pooled data were available for Small Dark Herons and Small White Herons without individual species identification. Data collected on each individual species were also pooled by feeding guilds.

Feeding guilds are defined as species with similar feeding methods and habitat utilization. These were defined as: Short-Legged Tactile Feeders (White Ibis, Glossy Ibis); Long-Legged Tactile Feeders (Wood Storks, Roseate Spoonbills); and Long-Legged Visual Feeders (Great Blue Herons, Great White Herons, Great Egrets). These guilds are somewhat arbitrary, as no two species are precisely the same in terms of habitat utilization or responses to hydrological or other environmental changes.

5.4 Mapping surface water patterns

During the monthly bird counts, each study cell was also qualitatively classified by the third observer in the plane into a surface water condition category based on the amount of the cell area covered by surface water. From 1985 through 1987, surface water conditions were classified into one of three categories: dry (absence of surface water), transitional, (a mixture of wet and dry) or wet (continuous surface water). From August, 1988 through May 1996, a five category classification scheme was used, consisting of dry/dry (absence of surface water and no groundwater visible in solution holes or ponds), dry/wet (absence of surface water but

groundwater present in solution holes or ponds), dry transitional (ground surface area mostly dry but small, scattered pools of surface water present and groundwater visible in solution holes or ponds), wet transitional (ground surface area mostly wet but small, scattered dry areas present and groundwater visible in solution holes or ponds) and wet (continuous surface water over the area). For this report, the 1988-96 five category metric was recoded to the 1985-87 three category metric by considering the dry/dry and dry/wet categories as dry, the dry/transitional as transitional, and the wet/transitional and wet categories as wet. The range of surface water condition patterns can be seen in Figure 14 where patterns for May of four years, 1990, 1991, 1995 and 1996 demonstrate how the dry season can end predominately wet, predominately dry or two different type of intermediate condition.

An index of water level condition was developed to characterize dry season surface water coverage as another categorization in the development of the HLC. The number of dry season months, averaged across all study years, that a grid cell remained in the "wet" category was defined as the average annual dry season condition for that cell. The index of dry season condition (Figure 15) for each cell was computed as the average annual dry season condition computed *using only those years for which water levels were determined to be normal* ('85-'88, '90-'93, '96).

5.5 Other variables of interest

For purposes of data analysis, the total study area was divided into several different geographical series and subareas or geographical strata, each series representing a different geographical classification of the study area. These geographic series and strata include: drainage basins or areas each with a different hydrological regime and/or water management scenario, major watersheds, major physiographic subregions, and major landscape or terrain patterns (Figures 3-11).

5.6 Notation

Let

n_{ij} be the count in the i th transect and the j th cell.

w is the width of the strip sampled.

A is the area of each cell, defined as $A = WL$ where

W is the distance along the North/South edge of the cell and

L is the distance along the East/West side of the cell.

N is the total number of birds in the total survey region.

N_{ij} is the number of birds in the ij th cell.

k_i is the number of cells in the i th transect, $i = 1, 2, \dots, r$.

$K = \sum_{i=1}^r k_i$ is the total number of cells.

Then

$D = \frac{N}{AK}$ is the total bird density,

DRAFT COPY

$D_{ij} = \frac{N_{ij}}{A}$ is the density in the ij th cell,

$D_i = \sum_{j=1}^{k_i} \frac{N_{ij}}{Ak_i}$ is the density in the j th transect, and

$P = \frac{w}{W}$ is the proportion of area sampled in each cell.

5.7 Precision and Accuracy of Counts

In any sampling design one needs to be concerned with the accuracy and precision of the population parameters estimated. In aerial transect sampling there are a number of potential sources of bias for estimators, some which can be controlled, some which can be estimated and estimates adjusted for and some sources of bias which cannot be taken into account. These will be discussed in the next section.

An evaluation of a survey of the type reported here needs to address three main issues. First is how precise are estimates of the measurements. This implies deriving some estimate of the variability associated with the transect/cell count.

The second issue deals with the accuracy of the counts. If the counts are precise then repeated measurements should average out to the true population density. There is much literature which addresses the question of bias in aerial surveys in general and systematic aerial transect samples in particular.

The final issue concerns the use of the transect/cell density estimate in the construction of a spatial distribution map. From this map can be had an analysis of temporal changes in the spatial distribution of the population as well as the examination of correlations between environmental/ecological parameters observed on the ground for each cell and that cells estimated density. Unfortunately, the question of the precision and accuracy of counts will affect the spatial distribution maps. The number of transects and number of cells used for each transect will also affect what information is available on the maps.

5.8 Precision

Precision is concerned with the repeatability of the survey method. Precise estimates have narrow confidence limits but the estimate itself may be biased. When discussing precision one needs to address which population parameter is of interest. We will discuss precision from three viewpoints, that of the estimate of total population size, that of the population size in a given transect, and that in a given transect/cell. We assume for the moment that the methodology produces accurate (unbiased) counts or equivalently that a mechanism exists that transforms biased counts into unbiased counts.

5.8.1 Method

For the SRF/WBS, by description of the study, the only true replication is that of the different transects. Cells within transects are really segments (split plots) and are usually not considered to be independent since counts from adjacent segments can be correlated. We will examine this assumption more closely in the subsequent analysis.

Analysis of precision requires the calculation of the variance of the population parameter estimates. For the total population estimate (or equivalently for the estimate of the total density of a specific class of bird) the variability among transect estimates is used in obtaining a precision

estimate. For transect densities, the variability among cells within the transect are used, adjusting for the possibility of dependency of neighboring cells within the transect. Since there is no replication of counts within the individual cells, an estimate based upon assumptions of the spatial distribution of birds within the cell as well as the nature of the aerial transect methodology can be used to derive a precision estimate.

5.8.2 Variance of the Total Density

An estimate of D_{ij} , the density in the j th cell of the i th transect is given as

$$\hat{D}_{ij} = \frac{n_{ij}}{wL}$$

The estimated count in the j th cell of the i th transect is

$$\hat{N}_{ij} = \frac{n_{ij}}{P}$$

The estimated total count for the i th transect is

$$\hat{N}_i = \sum_{j=1}^{k_i} \hat{N}_{ij} = \sum_{j=1}^{k_i} \hat{D}_{ij} A = \sum_{j=1}^{k_i} \frac{n_{ij}}{P}$$

The estimate for total density is given as

$$\hat{D} = \sum_{i=1}^r \frac{\hat{N}_i}{AK}$$

The estimate for total density in the i th transect is given as

$$\hat{D}_i = \frac{\hat{N}_i}{AK_i}$$

An estimate of the variance of D is given using the density estimates from the different transects and the total density estimate

$$\text{Var}(\hat{D}) = \sum_{i=1}^r \frac{(D_i - D)^2}{r-1}$$

This estimate assumes each transect line is a true replication, i.e., the r transects are as r random transects in the study area.

Alternate estimates of overall density and variances can be obtained using weighted averages based upon the fact that each transect is not the same length, that is does not have the same number of cells:

$$\hat{D}' = \sum_{i=1}^r \frac{k_i D_i}{K} \quad (1)$$

$$\text{Var}(\hat{D}') = \sum_{i=1}^r \frac{k_i (D_i - D')^2}{(r-1)K}$$

Other estimation procedures, such as jack knifing are available within this framework (Burnham *et al.* 1979).

5.8.3 Variance of Transect Density

For each transect, an estimate of variance may be obtained using the counts from the cells within that transect. Using the previous notation we define the variance for the i^{th} transect as

$$\text{Var}(\hat{D}_i) = 3S_{1i}^2 - \frac{S_{2i}^2}{(k_i - 3)b} \quad (2)$$

where

$$b = (k_i WL)^2,$$

$$S_{1i}^2 = \sum_{j=1}^{k_i} \frac{(n_{ij} - \bar{n}_i)^2}{k_i - 1}$$

and

$$S_{2i}^2 = \sum_{j=1}^{k_i-1} (n_{i,j+1} - n_{ij})^2.$$

This assumes that only neighboring cells are correlated, the correlation does not extend beyond the neighbor.

In **Appendix A** are given the estimates of the transect means, and the standard deviations calculated by using Equation (2) as well as the equation for a simple random sample.

$$V(D_i) = \frac{k_i}{k_i - 1} (A - B)$$

where

$$A = \sum_{j=1}^{k_i} \frac{n_{ij}^2}{k_i}$$

and

$$B = \sum_{j=1}^{k_i} \frac{n_{ij}}{k_i^2}$$

To indicate the degree of autocorrelation among cells within a transect, the Durbin-Watson statistic is also included in this table. The Durbin-Watson statistic is provided here to indicate whether there seems to exist dependence between neighboring cells or not. A significant D-W statistic suggests that some sort of dependency exists but does not tell one the nature of that dependency. The estimator given in Equation (2) assumes a first order autoregressive relationship exists in the data. This may or may not be the true situation.

From the tables in **Appendix A** we notice that the standard deviation estimate computed assuming correlated cells is usually larger than that computed assuming simple random sampling although the difference is usually fairly small. Many of the Durbin-Watson statistics are statistically significant at the 0.05 confidence level indicating that there is some indication of correlation among counts of neighboring cells. The standard error estimates are usually larger than the density average indicating that there is quite a bit of variability among the cell counts in

any one transect.

The overall density of the region changes from date to date. Confidence limits for the overall density estimates were computed and are given at the bottom of each table. Note that there are no statistically significant differences in dates across the sampling period since the 95% confidence intervals overlap. This implies that, at least from a statistical point of view, the total density did not change over the time period of the pilot program.

5.8.4 Variance of the Individual Cells

Since there is no replication of counts within a cell, no direct estimate of the precision of this estimate is available. Thus to obtain a precision estimate requires some assumptions.

The first approach to estimation of cell count precision requires the use of assumptions about the spatial distribution of individuals within the cell. Strip transect sampling, as we are discussing it here, implies that once we have chosen the center of the sampling strip, we will count all individuals within $w/2$ of either side of the center line. Consider the cell to be defined by the box in **Figure X** [I don't have this figure].

Assume the location of the center of the sample strip is chosen at random from the interval $[w/2, 1 - w/2]$. All individuals within the strip are observed (the circled 'x's), the others are not. The probability of observing an individual using this method does not depend upon its horizontal position within the cell (long direction) but only on its vertical position. Suppose the vertical position of each individual were plotted on the vertical end line. The final plot might look like that in **Figure Ya**. A histogram of these data might look like that in **Figure Yb** [I don't have these figures].

It is this distribution function (**Figure Yb**) which will determine whether our count is precise or not. For example, if the data are distributed uniformly across the width of the cell and the center line is chosen at random, then the distribution of the number of individuals counted in the sample interval will follow a binomial with parameters N_{ij} , the actual cell population, and $P = w/W$, the proportional width of the strip to the total cell width. Thus

$$\Pr(N_{ij} = n) = \binom{N_{ij}}{n} P^n \sigma^{N_{ij}-n} = \frac{N_{ij}!}{N_{ij}!(N_{ij}-n)!} P^n \sigma^{N_{ij}-n},$$

where $\sigma = 1 - P$ and

$$\binom{N_{ij}}{n}$$

is the number of set of size n from a total population of size N (read as ' N_{ij} choose n '). For a binomially distributed random variable we know that $E(n_{ij}) = N_{ij}P$, from which we get $N_{ij} = n_{ij} / P$, the estimate we used previously.

We also have that

$$\text{Var}(n_{ij}) = N_{ij}P(1-P). \quad (4)$$

Using

$$\text{Var}(\hat{N}_{ij}) = \frac{N_{ij}(1-P)}{P},$$

then

$$\text{Var}(\hat{N}_{ij}) = \frac{n_{ij}(1-P)}{P^2}$$

with

$$\text{Var}(\hat{D}_{ij}) = \frac{n_{ij}(1-P)}{(wLP)^2}.$$

If the distribution of the individuals is not uniformly distributed across the width of the cell then the estimate would be given as

$$\text{Var}(\hat{N}_{ij}) = \frac{an_{ij}(1-P)}{P^2} \quad (5)$$

where a is some unknown parameter ($a > 1$) which measures the level of departure from uniformity of the distribution. This parameter cannot be estimated using only one strip per cell.

Justification for use of the simple binomial estimate for cell density, Equation (4), is based upon two arguments. These arguments in turn depend upon the concepts of randomness in either bird behavior or randomness in the location of feeding sites. Although at this point the following speculation cannot be validated, we offer these arguments as possible mechanisms for sufficient randomness.

Although one does not know the exact location of all birds within the target cell, two possible mechanisms exist which would render bird locations somewhat random with respect to their vertical position (i.e., north-south location) within the cell. During enumeration a bird is either flying or stationary. If flying, the location of the bird will depend upon the time at which it is spotted, which given all the events necessary for this to occur could be considered a random event. If stationary, the bird may be feeding, most probably, or not feeding, less probable, depending on factors such as time of day, etc. Given that transects are oriented, in a sense, across the general direction of flow of water in the slough, likely feeding sites will be oriented at random to the north-south axis of the cell. Groups of birds will have correlated locations with regards to north south orientation, but when a number of groups are located within the cell the distribution of vertical locations of individuals within the cell will be close to uniform.

If one does not feel that the birds or habitat are cooperative, neighboring cells along a transect could be pooled to assure the vertical uniformity. (The process which guarantees this operates similar to the central limit theorem.) For this combination not to work would require some mechanism for having more birds in a certain section of cell (or string of cells) than in other parts. This would in all probability be noticed by the enumerators and adjustments to the survey made to take this into account. If not taken into account it would result in not only imprecise estimates but biased estimates. All other assumptions for nonuniformity, in the author's opinion, assume that nature is perverse, conspiring not to give one a decent chance at a decent estimate. We are open to enlightenment on this topic.

5.8.5 Alternative Cell Variance Estimation Procedures

Borrowing from methods for analyzing remotely sensed data, one could use a smoothing type of estimator based upon a spatial averaging of neighboring cells. A group (window) of neighboring cells are used to estimate the density at the 'center' cell and the variance of this estimate computed using the differences in counts among the neighbors used. Thus, for example, one could use sets of five contiguous cells within a transect and estimate D_{ij} using Equation (2). That is, let the estimate of density, D_{ij} be given by

$$\hat{D}_{ij} = \frac{(n_{i,j-2} + n_{i,j-1} + n_{ij} + n_{i,j+1} + n_{i,j+2})}{5wL}$$

and

$$\text{Var}(\hat{D}_{ij}) = \frac{3S_{i1}^2 - S_{i2}^2}{50^2 wL},$$

[Ken—is this right? Original was: $\hat{\text{Var}}(\hat{D}_{ij}) = \{3[S_{i1}^2] - [S_{i2}^2]\} / (50^2 wL)$]

where

$$S_{i1}^2 = \sum_{l=1}^5 \frac{(n_{i,j+l-3} - n_{ij})^2}{4}$$

and

$$S_{i2}^2 = \sum_{l=1}^5 (n_{i,j+l-2} - n_{i,j+l-3})^2.$$

This provides an estimate for the ij th cell using the cells on either side of it. This also assumes correlation between neighbor cells as before. A smoothing of the observed counts (a running average) is performed which tends to reduce differences between neighbor cells. If indeed there is correlation between cells, this method will in one pass provide some smoothed estimates of density and an associated measure of variability. If the cells are actually independent, then we will have lost some ability to discriminate between cells of different density and in addition have an inflated estimate of the precision variance.

Estimates of variance for the cells on the ends of the transects could be obtained by weighting the end cells more heavily in the smoothing function. Properties of the variance for these cells are not known.

To determine whether this model would be appropriate would require examining the auto-correlogram or partial auto-correlogram of each transect for a time series structure. This may not be very useful if the counts (or log counts) in each cell are quite low. This has not been done for these data since little information would be gained.

A second alternative approach would be to figure into the calculations of the last method the information from neighboring transects. Thus, for example, the density for the ij th cell would be

$$\hat{D}_{ij} = \sum_{l=1}^3 \sum_{m=l}^3 \frac{n_{i+l-2,j+m-2}}{9wL}.$$

[Ken—is this right?]

Thus D_{ij} is a spatial smoothing (a low pass filter) of the eight (8) points around ij and the value in n_{ij} . The variances of D_{ij} consists of two components,

σ_b^2 is the among transects variability, and

σ_w^2 is the between cells within transects variability.

This variance would still be estimated as

DRAFT COPY

$$\hat{\text{Var}}(\hat{D}_{ij}) = \sum_l \sum_m \frac{n_{i+l-2, j+m-2} - n_{ij}^2}{8wL^2}$$

Note that this estimate takes into account neighboring transects but does not account for correlation. Again we get a smoothing of the density estimates in order to obtain an estimate of precision for the cell count.

5.8.6 Summary

The results of this section can be summarized as follows.

1. Total population density precision should be estimated using Equation (1).
2. Total transect density precision should be estimated using Equation (2) rather than Equation (3) [Ken—which is equation 3?]. The results of the Durbin-Watson statistics indicate the presence of some correlation between neighboring cell counts. Equation (2) may not be the best estimate for these data but represents an improvement over the simple estimate.
3. Cell density precision should be estimated using Equation (4). Some effort should be made in the future to validate the underlying assumptions of this method. The spatial windowing method of estimation smooths the spatial (map) distribution, with resultant loss of resolution and possibly loss of correlation with habitat environmental parameters.

5.8.7 Future Work

1. Attempts need to be made to validate the assumptions of the cell density precision estimator or find a more acceptable estimator. Initially, a Monte Carlo computer simulation will be performed to examine the value of a in Equation (5) as a function of different nonuniform distributions.
2. Attempts need to be made to prove that uniformity could be achieved by using a long enough cell, regardless of the underlying spatial distribution.
3. Attempts need to be made to determine how long the cell should be for the uniformity condition to be essentially valid.
4. The reasons for correlations among neighboring cells within a transect need to be examined.

5.9 Accuracy

Accuracy is concerned with the closeness to which cell count predicted (or calculated density) is to the true value in the cell. We say an estimated value is biased if its expected value is not the true population value we seek. For aerial survey data obtained via systematic strip sampling, a number of sources of bias have been discussed in the literature.

M. Norton-Griffiths (1975) discusses the problem of bias in aerial transect sampling. Briefly these are listed below.

1. Bias from sample design. Designing the study with built-in biases by avoiding the use of sampling techniques of stratification and random sampling where appropriate.
2. Bias from definition of the population under study. Bias can enter the study by not carefully defining the population under study or by sampling the wrong population or part of the population. Factors such as time of year, migration, study area boundaries are important at the design stage whereas errors in navigation, plane speed, height and strip width variability and incomplete counting during the survey are of primary importance. Using the wrong

values to estimate density from counts can produce biases at the analysis stage.

5.9.1 Counting bias

This is considered the most important source of bias for the aerial transect sampling method. This includes differences in observability between on plane observers due to differences in experience. The height, speed and strip width used in the survey affect the ability of the observers to count birds. Different habitats also have different sightability characteristics.

Although all sources of bias cannot be eliminated, many can be controlled via proper design, adherence to that design and careful analysis.

This pilot study was designed in an attempt to eliminate the obvious biases. Transects were run east to west, across the direction of flow of water in the slough, in an attempt to avoid systematic bias due to a transect being always in a wet or dry 'run'. Each cell thus contains both wet and dry areas. Habitat was restricted to the slough areas for this pilot, avoiding the mangrove habitat to the west and south. Areas were stratified into the major sloughs and the two containment areas to the north of the park.

All flights were flown at one fixed height, speed and strip width, with a separate study performed to examine the effect of these factors on sightability. A one km square cell was chosen in order to have the highest resolution possible for spatial distribution analysis as well as a large sample size for estimation of transect and total population density.

This section will concentrate on the analysis of the data on counts from the separate study of height, speed, and strip width factors along with a preliminary analysis of time of day effects and observer differences.

5.9.2 Methods

For the SRF pilot study a separate study (experiment) was performed to gather data necessary to obtain correction equations for the observed counts. This approach was suggested by Caughley, *et al.* (1976) who first discussed a quantitative method for bias correction of aerial transect counts. His approach was to use a balanced factorial experiment with Analysis of Variance and Regression techniques to explore the effects of several factors, in particular, speed of the plane, height of the plane and width of the strip scanned, on the accuracy of transect counts. The end result of this exploratory analysis is an equation in some of these factors which 1) indicates the expected proportion of individuals actually in the strip which were observed and 2) can be used to obtain a correction term for removing the effects of this bias from the density estimates.

The foundation of this analysis is the balanced factorial experiment. In particular, three levels each of height, width and speed are chosen at which counts in a fairly homogeneous region are made. The three levels of each factor are arrayed in a factorial design, that is, each possible combination of factors are used. Flights were made at each of the factor combinations and counts were observed. The design was repeated for a number of days in order to get a good average of the effect of each factor.

A number of other factors such as time of day the observation is made, and date since beginning of experiment, were also recorded. This preliminary analysis will not address these factors but their effect will be examined in a later analysis.

The preliminary analysis consist of examination of the analysis of variance tables for counts of total population as well as that for the subpopulations listed below.

Total birds-All size classes-Excluding flocks

Small dark birds-Excluding flocks

DRAFT COPY

Large dark birds-Excluding flocks

All dark birds-Excluding flocks

Small white birds-Excluding flocks

Large white birds-Excluding flocks

All white birds-Excluding flocks

5.9.3 *Analysis of Variance*

The analysis of these data were performed upon the square root of the count plus one as suggested by Caughley *et al.* (1976). This transformation is needed to provide some semblance of homogeneity of variance to the counts. and is recommended statistical practice.

The analysis of variance (ANOVA) allows us to identify those factors or sources of variation which are statistically nonzero. It is the first step in the development of sightability curves. It tells us which factors should be considered in the subsequent regression analysis.

The analysis of variance tables for the data for each of the populations listed above are given in **Appendix B**. Significant effects (factors) are identified by very small (< 0.05) values in the column labeled $PR > F$. The analysis was performed using the ANOVA procedure in the SAS package (SAS Institute Inc. 1982). A copy of the program used is included in **Appendix B**. In all cases the analysis was also run on the untransformed data with no change observed in the analysis conclusions.

The development of bias correction equations is done through the use of regression techniques, in particular stepwise multiple regression. Again following the suggestions of Caughley *et al.* (1976), stepwise multiple regression was used to find the 'best' equation of the factors and their interaction terms which described the actual counts. Note that the analysis is performed on the actual counts at this point not on the transformed data. The regression equations finally chosen, after examination of the ANOVAs, stepwise regressions and R-square values of all possible models are given in **Appendix C, Table C.1**.

5.9.4 *Summary (ANOVA)*

From the analysis of variance, a few general observations may be made.

In all cases, there was no observer effect. This implies that the observers were essentially observing the same population with roughly the same level of bias. This is not too unexpected with experienced observers.

There were no significant effects for the small dark bird subpopulation. This could be the actual situation or an artifact of the low number of such birds observed in the study. The low number of small dark birds could have been the result of observer bias or actual low population density. Whatever the situation, our experiment was unable to come to a bias adjustment function for this subpopulation.

In all other populations, the transect strip width factor was involved. In most situations the factor was involved as part of an interaction with one of the other factors. By far, strip width is the major factor in bias.

Height of plane was the next most important factor usually involved in the bias correction analysis.

All ANOVA *r*-squared values were very small indicating that much variability in the data has not been accounted for by the factors examined. We note that while Caughley advocated this

technique, his published works show the same low levels of explained variation as seen here, with no comment.

Finally, speed of the plane seemed to only affect sightability for the different size class population.

5.9.5 Regression

The regression equations given in **Table C.1** were developed using multiple regression methodology, tempered by the results of the analysis of variance. Parameter estimates and their standard errors for the final models for each size class are presented. Unbiased estimates of the true density, assuming the fitted model is a valid representation of the sightability curve for this habitat, time period, are given as the intercept term of the curve. Thus, for example, for large dark birds an 'unbiased' estimate of the actual density is 1.836 per transect area, with standard error of this estimate given as 0.332 individuals per transect area.

The method goes on to compute the equations for predicting the percent counted as a combination of the factors which affect bias. These equations are obtained by simply multiplying through each regression equation by the quantity $100/(\text{intercept term})$. These equations are given in **Table C.2**. The quantity denoted as 'correction term' is simply the reciprocal of the intercept times one hundred percent.

Plots of the predicted proportion observed equations are given in **Appendix D**. To illustrate their use consider the plot of large dark birds versus height of plane by strip width (the first plot). Since the proportion observed is a function of strip width and height, we have a three dimensional picture to display. The two dimensional plots given allow one to choose any value of height and examine the expected proportion for any of 9 different widths. Thus for a flight at 200 feet, one could expect to observe roughly 50% of individuals in a strip 70 meters wide but only 15% in a strip 220 meters wide. Note that this particular plot will not converge to 100% unless both height and width are zero.

From the plots of the proportion observed equation we can examine how different survey designs will effect the proportion observed. We will point out here that there is a great deal of variability which is not accounted for by each of these equations. This is evidenced by the very low *r*-squared values associated with each equation. This implies that the prediction capability of these equations is very, very limited.

5.9.6 Summary

The analysis of variance indicates that some of the factors examined, specifically strip width and plane height, showed significant differences in counts at different levels and widths. This would seem to imply that adjustments in observed counts should be made depending on the final choice of these factors levels in the next study.

Due to the large amount of variability not explained we can conclude that the effect of these factors, while statistically significant, may be minor when considered in the light of the total picture. Other factors, such as day to day movement of birds, micro habitat effects and other environmental conditions could play a far greater role in explaining cell to cell and day to day variability.

The regression models assumed, i.e., polynomial with interactions, may not be appropriate for these data. This is illustrated when one examines the individual estimates of large versus small white birds, 3.43 and 0.736 individuals per unit area respectively, and compare these to that for all white birds, 11.135. These estimates may be unbiased, if one is willing to assume the linear type models used here, but they are certainly not precise.

5.9.7 *Future Work*

Future work on the accuracy of the counts lies in two directions. First, other sources of variation need to be examined. Specifically, habitat effects, expected to be a significant source of sightability differences, was not examined in this study. A study, similar to this one, which incorporates habitat, height and strip width at a minimum should be run.

Another factor which was not examined was the effect of density on sightability. The more birds in the sample area the harder it is to count them all, leading to some sort of threshold effect. If possible the study should be done during high bird densities. The high counts obtained will be helpful in satisfying some of the ANOVA and regression assumptions.

The second direction for future work lies in the development of better analysis methodology. Since the data we work with are counts, the first place to look for methodology would be the recent literature on the analysis of categorical and count data (see Nelder and Wedderburn (1972) and Macullah [McCullagh?] and Nelder (1983)).

5.10 *Spatial Distribution and Grid Cell Resolution*

Systematic aerial sampling, of the type discussed here, has the advantage over random sampling in that a distribution map, in addition to a population estimate, can be easily obtained from the data. This can be important if one is interested in the distribution of animals relative to habitat conditions. In this section we will discuss the relationship between the size of the cell represented by our sample and the resolution of the distribution map obtained from that sampling. Comments on habitat / population distribution analysis will be given in the next section.

5.10.1 *Methods*

The question of grid cell size is similar to that of choice of quadrat size in ecological studies. In the SRF/WBS, one side of the quadrat is fixed, that is the width of the sample strip. We then have a choice of the length and width of the cell through which we fly the strip. If the cell through which we fly the strip is fixed at some value W , then the proportion (P) of the area covered by the sample strip will be the same regardless of the length of the cell. In Section 1, we showed that the expected variance of the density estimate in a cell is a function of the proportion P , not of the length. Thus decreasing P will decrease the confidence we have in our cell density estimates. Adding length to the cell just assures that the distribution of individuals across the width of the cell will be more uniform, hence increasing our confidence in the variance estimate.

Decreasing P corresponds to flying fewer transects across the region at any given time. Decreasing P also means that there will be fewer transect/cell combinations available for estimating the population density as well as creating a distribution map. Thus we need to access the effect of a reduction in P on these values.

The data from the pilot study was obtained on a 1km x 1km grid spacing using a P of about 0.35. From these data we can examine what we would have gotten if certain other grid spacings had been used. Thus for example, if we use every other transect, results from a grid spacing of 2km x 1km can be examined. Adding cells within a transect can make a grid size of 2km x 2km. Similarly, we can examine sizes of 3km x 1km, 4km x 1km, etc. The length of the cells can also be changed at no extra cost simply by combining the 1km long cells.

This method of analysis is called resampling. We resample the data we have only assuming a part of these data can be used at any one time. The results of this analysis will demonstrate the range of values for population parameters and spatial statistics under less intensive sampling.

DRAFT COPY

5.10.2 Explanation of the Resampling Scheme

There are 52 1km wide transects in the pilot study. If we chose every other transect, in effect assuming a 2km wide cell, then there will be only 26 2km wide transects. For 3km wide cells there will be only 17 transects and for 4km wide cells there will be only 13 transects.

Using the data from the pilot study it is possible to define w^r realizations of a study using r w km wide transects. To see this, suppose there were only six (6) transects and we wish to consider using only three at a time from the three sets of pairs. Then we have from the following table that there are $2^3 = 8$ populations using $w = 2, r = 3$.

Original Sample	Populations of 3 transects from adjacent pairs						Populations of 2 transects from adjacent triples							
1	1	1	1	1				1	1	1				
2					2	2	2	2			2	2	2	
3	3	3			3	3						3	3	3
4			4	4			4	4	4		4		4	
5	5		5		5		5			5		5		5
6		6		6		6		6		6		6		6

Each of these w^r samples would produce a different value for the overall density as well as a different distribution map.

5.10.3 Statistics Examined

Due to the computer expense of this resampling exercise and the time limitations on this report, analysis was limited to three statistics. Specifically, the overall average density from a specific (re)sample for the area was examined along with the variance among (re)sample density estimates. This was done to explore the effect of different grid cell sizes on the estimate of the overall population density estimate (objective one of the survey). One might expect the average density estimate to change little as sampling intensity decreased for reasonable grid sizes. One would also expect the variance of the average density estimates to increase as grid size increased.

Quantifying the effect of grid cell size on spatial (map) distribution is a much more difficult process. Since a map is inherently visual in its total information content, it is difficult to totally describe the map in terms of statistics. Concepts such as degree of aggregation cannot be adequately summarized by one statistic. As a first attempt to examine the effect of grid cell size on spatial distribution, we used the variance/mean ratio of cell counts for the resampling study. This ratio has been used as a measure of spatial aggregation (see Pielou, 1977). A ratio close to one indicates a random distribution of counts, larger than one indicates an aggregated population and less than one indicates a uniform distribution. The distribution of these ratios from a number of resampled populations for different cell sizes could indicate whether a change in cell size could change our perspective of the spatial distribution of counts.

5.10.4 Results

By resampling transect densities from the original data we can generate a plot of the statistical sampling distribution of the overall density estimate. Frequency plots of these sampling

distributions are given in **Appendix E**, along with plots of the variance of the overall density estimates. Means and variances of these distributions are also given in **Appendix E**.

The plots of the means and standard error estimates of these means are given in **Appendix F**. By examining the change in the standard error estimates as the resolution changes we can get some feel for how the overall picture of the data would change. In most cases the 2km resolution (that is 2km wide by 1km long) provides the smallest standard error. For February there is a small increase in std. err. as we go from 1km x 1km to 2km x 1km, but a much larger increase as we proceed to 3km x 1km cells.

Combining cells to create new cell sizes within a transect will have no effect of the overall density estimate. What will happen is that the resolution of the distribution map will be less. The effect of reducing the number of cells can be observed by employing a measure of spatial pattern or aggregation. As the size of the cell is successively increased, plots of the aggregation statistic will indicate our changing perception of the nature of the population distribution. The statistic we use here is the variance/mean ratio of all cells. A plot of these values for different cells sizes is given in **Appendix F**.

The frequency plots of **Appendix F** illustrate the effect of reducing the number of transects on total population mean density and the variance/mean ratio. Specifically, we see that the distribution of expected mean density will change very little as the resolution decreases. As seen in the plots, the distribution spreads out indicating an increased range of possible mean values. The trimodal nature of some of the plots illustrates the effect certain high density transects can have on the distribution. When a high density transect is included in the sample we can expect a dramatic increase in the overall density estimate as well as the overall variance.

The frequency plots of the variance/mean ratio illustrate the changing nature of the spatial distribution of bird counts as the sampling resolution decreases. Note that as with the average density distributions, the ratio distributions spread out as resolution decreases. Thus sampling less can result in a 'picture' of bird distributions which shows more clustering than we would have seen in a more intensive sampling. Again the effect of high density cells can be seen in the trimodal nature of some of the plots.

If one is forced to reduce the amount of sampling in future studies it seems wise, based upon these results, to sample to the limit of the available economics. Any reduction in sampling will result in a less clear picture of bird distribution.

The 2km x 1km grid cell seems to provide the next best sample plan for production of a distribution map. Since this represents half of the flying time used in the pilot study, the associated reduction in costs are reduced proportionately.

Conclusions on the effect of increased grid cell length, say to cells of size 2km x 2km, 2km x 3km, 3km x 3km, etc., will need to await further computer runs and some discussion of how cells should be combined. The relatively minor cost involved in providing the data on the 1km cell length is a powerful argument for collection of data using that width.

5.10.5 Future Work

It should be obvious that this is only a partial answer to the question of grid cell resolution. Much more needs to be done on this topic, from both a theoretical and practical viewpoint.

We suggest that the next level of analysis of this problem be to utilize other measures of spatial distribution in the resampling study. In addition, other approaches, such as the spatial averaging (moving window averaging) might provide further insight.

DRAFT COPY

6 Surface Water Condition: Temporal Patterns

Each transect/cell on a survey date was classified as "dry", "transitional" or "wet". The proportion of the 1075 cells allocated to each water condition class forms a simplex. Simplex data can be plotted on a triangle, where the nodes of the triangle represent 100% of the cells in one of the three classes. The water condition proportions are presented in **Figure 22** with line segments connecting temporal neighbors and arrows indicating the direction of time. The general pattern is one of a cycle. Starting in August the study area was predominately wet. There was a tendency for the percentage of dry cells to increase from December to May then decrease dramatically in August. At the same time, the percentage of wet cells decreased from December to May with a subsequent dramatic increase in August. The path from predominately wet to predominately dry was usually through a predominately transitional stage. The fact that the path is not a simple open form (circle) but is "filled" suggests that there was a lot of variation among years in the actual cycle.

The proportion of transitional cells and wet cells that did get drier (P_D), the proportion of dry cells and transitional cells that did get wetter (P_W), and a 'change potential index' (CPI) were used to examine the direction of change in surface water condition. All of these addressed the question of whether cells got drier or wetter as the year progressed. Computational details for these measures are given in Appendix C. CPI values range from -1 to 1, with positive values indicating a drying trend and negative values indicating a wetting trend. Temporal patterns in CPI, P_D and P_W were examined graphically.

Time plots of the proportion of cells falling into the three hydropattern categories were examined for trends. The stability of the drying pattern in each cell over the dry season was assessed by examining the change in hydropattern classifications for each of the intervals: December to January, January to February, February to March, March to April, April to May, May to August, and August to December. No change in condition class between consecutive months produced a 'stall' classification for that grid cell. A "reversal" in the drying pattern was recorded when the second of two such consecutive values represented a wetter category than the first. "Progressive drying" was recorded when the second value represented a drier category than the first value.

The stability of the surface water condition pattern in each cell over the dry season was assessed by examining the change in water condition classification for neighboring months: December to January, January to February, February to March, March to April, April to May, May to August, and August to December. No change in condition class between consecutive months indicated a 'stall' classification for that grid cell. A "reversal" classification was recorded when the second of two such consecutive values represented a wetter category than the first. "Progressive drying" was recorded when the second value represented a drier category than the first value.

A weighted kappa statistic (Agresti, 1984, Cohen, 1960) was used to index persistence of a surface water condition over a two month interval for the whole study area. This index ranges between values of 0 and 1, with a value of 0 indicating persistence consistent with no association in category assignment between the two months, and a value of 1 indicating perfect persistence of pattern, i.e., no cells changed category, a very non-random situation. As the kappa index goes to zero, fewer cells have the same category in the second date than they had in the first date. The occasional negative weighted kappa value indicates a change pattern displaying no consistent pattern but which has less agreement than would be expected by chance. The weights are used to account for the fact that the probability of going from say "wet" to "transitional" may not be the same as the probability of going from "wet" to "dry".

A graph of the weighted kappa statistics (**Figure 23a**) showed a consistent year-to-year pattern in water condition persistence. The level of agreement in consecutive classifications for transitions:

Dec-Jan, Jan-Feb, Feb-Mar, Mar-Apr, Apr-May is constant for most years, with less agreement observed for the May to August transition, and only moderate agreement for the August to December transition.

A graph of the drying proportion, P_D shows no clear dry season trend, but a trend could be hidden by the large amount of variation observed. P_D drops from May to August then rebounds from August to December. The wetting proportion, P_W , shows the opposite trend. The change potential index, CPI, does not demonstrate a trend during the dry season but there is a decrease from May to August and subsequent increase from August to December except for 1994.

7 Bird Counts: Temporal Patterns

Bird populations in the Everglades are dynamic, ever-changing. One goal of the overall study is to examine the temporal changes in population and attempt to correlate some of this movement to changes in desirable habitat. In this section we discuss how the temporal changes in the population could be analyzed and quantified. The next section discusses how these changes might be correlated with changes in the environment.

Total observed wading birds summed over all 93 sampling dates are presented in Table 8 along with the distribution of these counts to feeding guilds and individual species. These numbers are misleading since they do not take into account the spatial and temporal variability in estimated counts, but are presented here as an indication of the total sampling yield for the effort presented.

7.1 Temporal autocorrelation

Analysis of the temporal aspects of population dynamics can be performed in a number of ways. A first approach would be to examine the correlations between counts at time one and those at time two. A strong positive correlation would imply little or no change during the intersampling time period. Zero or negative correlation would imply change in the spatial distribution and would indicate the need for further study. This follow-up analysis could consider changes in the statistics of spatial aggregation, spatial autocorrelation and overall density. To illustrate this we will examine the correlations between the cell counts for the five samples of the pilot study. Pearson product moment correlation coefficients and the Spearman rank order correlation coefficients are given in **Appendix G**.

7.1.1 Summary (Total Birds)

There seems to be very little correlation between total bird cell counts among the different dates. The highest correlations are about 0.11.

The AM versus PM counts for 26 April also show little correlation ($r = 0.0125$, with significance probability $p = 0.833$).

7.1.2 Summary (By Size and Color Class)

With the exception of Wood Storks, there is no correlation between dates over the whole study region.

A correlation of 0.6 was noted between the 29 February and 12 March counts for Wood Stork. Since this is the only high correlation in the complete set, one should not place too much emphasis on this finding.

Due to the large number of cells having birds, many of the correlations have low (< 0.01) significance levels. This indicates that these correlations are statistically different from zero. This is relatively unimportant, it is the magnitude of the correlation as well as its significance which is

examined for effect.

7.1.3 *Summary (Size and Color Class by Basin)*

For all birds: Northern Taylor Slough, East Slough and Shark Slough all show correlations which are greater than 0.1. Highest correlations were recorded for Northern Taylor Slough.

By Species: There are a scattering of middle range correlations (0.4–0.7) in the by species by basin tables, too many to identify individually.

7.1.4 *Future Work*

Examine the distribution maps to see if transect/cells could be combined into regions exhibiting high or low densities over time. This stratification could occur within a basin or cross basin boundaries.

Explore the effect of spatial averaging, discussed in the precision section, on temporal correlations.

It is possible to construct various maps which might illuminate temporal changes missed by the statistics. One possibility would be to simply subtract one date's counts from a subsequent date's counts and plot the resulting value. Although not very statistical such maps contain a large amount of information.

8 Bird Counts: Spatial Patterns

8.1 *Spatial autocorrelation within species*

This issue will be addressed in the next section entitled Bird Density Estimation.

8.2 *Associations Among Guild/Species*

Species associations and feeding guilds associations were examined using the Jaccard index (Pielou, 1984) calculated for each pair of species for each date. The Jaccard index was chosen over others such as Dice or Ochia because of its ease of interpretation for these data. An index value of 0 indicates no association while a value of 1 indicates perfect association. The index only measures presence/absence and does not compare densities of the birds. If two species were not seen at all during an entire survey the indices were not computed. The index was plotted over time to look for trends. A test for strongly associated species pairs was performed (Ludwig and Reynolds, 1992 [1988?]).

8.3 *Spatial patterns 1: by basin*

Tables of the breakdown of counts of occupied cells and numbers of birds per cell were produced by basin. These tables are given in **Appendix H**. The tables should be self explanatory, so no interpretation is provided.

8.4 *Spatial patterns 2: by cell*

Spatial pattern by cell are best viewed as maps. Maps of both cell bird count and cell water condition accompany this report: we refer the reader to those.

9 Bird Density Estimation

Two approaches to estimating bird densities and computing corresponding confidence intervals were examined. One approach used the ratio method of population estimation (Jolly, 1969; Krebs, 1989) that considers each transect to be an independent sampling unit. With this approach, only 35 unequal length sampling units were available for estimating the standard error of the estimator, resulting in wide confidence intervals. Count data in the cells along a transect are considered subsamples and the spatial arrangement of counts in cells along a transect is essentially disregarded. This method also disregards the discrete nature of count data.

Estimating bird densities for the entire study area using spatial location of cells requires that correlation between observations be accounted for. Failure to incorporate correlations can result in inflated (positive correlation) or deflated (negative correlation) estimates of the standard errors of the density estimates and hence produce incorrect confidence intervals. In this study, it was unlikely that all cells on a transect would be equally correlated for two main reasons: cells on the east and west edge of a transect can be up to 80 km apart and wading birds tend to aggregate when foraging.

The second approach, and the one reported here, incorporates the spatial information obtained in the cell counts and models the correlation between counts of neighboring cells. Using Generalized Linear Models (GLM) and Generalized Linear Mixed Models (GLMM) methodology (Appendix A), density estimates and associated confidence intervals were computed for the whole study area. These estimates were then converted to an overall abundance estimate. Note that these estimates were not adjusted for visibility biases and hence are very likely underestimates of the true area abundance. This approach was considered acceptable since the objective of this study was more to compare estimates across time than to accurately estimate abundance. It is assumed that the effects of visibility bias in the aerial transect methodology are consistent throughout the time period of the study, hence making comparisons across time appropriate.

Because of the long computational times needed to estimate model parameters assuming full spatial correlation, the full model was fit for a subset of dates, selected across the whole time span of the study. Based on these results a simpler correlation model was used for density estimation for all dates and for all species. Six correlation structures were examined.

Independence: No correlation is assumed between cells or transects. This method was equivalent to calculating the sample mean count and standard deviation using all 1075 cells as independent observations.

Compound symmetry: All cell observations along a transect are assumed correlated, the correlation is the same for all transects, and cells in different transects are not correlated.

AR(1): Cell observations along a transect exhibit the correlation structure of an autoregressive process. Thus, if two neighboring cells have a correlation of ρ , cells that are 'one cell apart' have a correlation of ρ^2 , cells that are 'two cells apart' have a correlation of ρ^3 , etc. Cells in adjacent transects are assumed uncorrelated.

AR(1) with random transects: The same correlation structure as above is assumed. In addition, transects total counts are themselves assumed random. This model accounts for transect-to-transect variability.

Spherical, no nugget: Cell observations along and between transects are assumed correlated. In geostatistics this is referred to as the isotropic correlation model since the correlation between two cells does not depend on compass direction, only distance. (see Littell *et al.* 1996).

DRAFT COPY

Spherical, with nugget effect: Cell observation correlations follow the same structure as the no nugget model, but allows for a nugget effect (see Littell *et al.* 1996).

Table 6 gives more details about the models.

The models listed above are of quite different structures and are not nested, hence no statistical test is available for assessing whether one model fits better than another. Instead, models were compared using Schwartz's Bayesian Criterion (BIC) for four dates. The BIC allows comparison of models after adjusting for the differences in the number of parameters estimated by each model. Larger values of BIC indicate better model fit. To better meet the assumption of normally distributed errors, counts were transformed to $\log_e(\text{count} + 1)$. Table 7 give the BIC values.

The spherical model with nugget effect had the largest number of parameters, most complex correlation structure and highest BIC values; however, it took hours of computational time to estimate model parameters for just one date for one species. This model was simply too complex for screening 93 dates for 13 species/guild combinations. A large jump is seen in the BIC when comparing the AR(1) model to the independence model. The jump in BIC was much smaller when comparing AR(1) to either of the spherical models. Also, because counts were made in sequential cells along aerial transects, one would expect there to be more correlation between cells within a transect than between cells on different transects. For these reasons, the AR(1) model was chosen as the base model for estimating bird densities.

Statistical inference with mixed linear models requires the assumption of normally distributed errors. Usually for count data, a $\log_e$ transformation is applied and log counts are assumed to be normal. However, the natural log transformation can distort the correlation relationships among counts and produce unacceptable negative variance estimates. Using a generalized linear mixed model (GLMM) it was possible to assume the more appropriate discrete exponential family of distributions (e.g., Poisson) for the count data while also incorporating the Autoregressive AR(1) structure for expected counts along each transect. A property of the Poisson distribution is that the variance of cell counts is a positive multiple of its expected value. For spatially distributed points, if the point pattern is random, cell counts should be Poisson distributed which makes the mean and variance of the counts the same. By allowing the variance of counts to be a multiplier of the expected count, a more flexible and general model is used. This extra variability assumption is examined by estimating the overdispersion parameter and testing whether it is significantly different from one. An overdispersion parameter equal to one indicates that counts follow the Poisson distribution, and is indicative of a random spatial pattern. An overdispersion parameter much greater than one indicates that the counts are aggregated or clustered. Significant overdispersion will affect the estimates of the standard errors and associated test statistics. The software used to fit this GLMM(Poisson) model (GLIMMIX macro in SAS, Littell, *et al.*, 1996) automatically adjusts the standard errors and test statistics to account for overdispersion. Using the results of this model fitting, a program was constructed to estimate the density (birds/km²) and associated 95% confidence interval.

Using the GLMM(Poisson w/ Autoregressive AR(1)) models discussed in the methods section, estimated abundance and associated 95% confidence intervals were computed. Time plots of the estimates and confidence intervals are presented in **Figures 17 to 21**. An indication of the degree of temporal variability in these estimates can be obtained from Table 9.

10 Bird Density Prediction: Linear Models and Land Element Themes

To examine the impact of changes in the environment on biological populations, one needs to

DRAFT COPY

know something about the relationship between these two. Much of the information on environment (habitat)/population relationships is qualitative, developed over long years of intensive observation of individuals in small (micro) environmental systems. The goal of this study is to examine the macro environment and its relationship to large area changes in population distribution. We want to quantify the relationships present in order to be able to explore the effect of different environment management schemes on the population.

The immediate concern of this study is the development of a quantitative model for the effect of habitat and water level on bird spatial distribution. The hypothesis is that a change in water level or habitat will necessarily produce a change in the density of birds in the impacted region or a rearrangement of birds into a different spatial pattern (random distribution to aggregated distribution say or vice versa).

An attempt was made to describe the variation in wading bird cell counts by using the land element characteristics described above: soil type, ground surface features, vegetation, surface water condition, surface water condition change and surface water (dry season) hydroperiod, in a generalized mixed-linear model analysis. Because of the size of the data set and abundance of hypotheses of interest, the analysis was performed in three phases progressing from simple models to complex models.

Phase 1: Wading bird/landscape relationships were investigated by linear regression of $\log_e(\text{count} + 1)$ as the dependent variable using the environmental parameters identified above as linear predictors in the model. The goal of this analysis was to identify where associations might be strong. Separate analyses were done for each sampling date, cells on the same transect were assumed autocorrelated, and transect-to-transect differences were assumed to be random effects.

Phase 2: The best of the models from Phase 1 were refit using generalized linear models with independent Poisson distributed error structures. This model is the most appropriate for count data of this type, and provides the clearest picture of associations.

Phase 3: A limited multivariate model approach was attempted in which correlations among bird counts were incorporated along with the correlations among environmental variables and time.

10.1 Decomposition of total variability in bird counts and importance of environmental correlates

The major objective of the data analysis was the search for temporal trends and spatial patterns in wading bird counts. A generalized linear modelling approach was used to assess the importance of long-term temporal trends and inter-year cycles in explaining total variability in wading bird counts. The residuals from these models represented unexplained temporal and spatial variability. The importance of land element theme classifications (Table 3 and 4) and landscape classifications (Table 5) in describing average spatial pattern was further examined by including these variables one-at-a-time as covariates in the log-linear models. Because these environmental correlates were only available as snapshot images for one period of time, it was not expected that they would explain a very significant fraction of the total temporal/spatial variation in counts.

Surface water condition was also used as a covariate in the model to examine its importance in explaining both temporal trends and spatial pattern of counts. Because surface water condition was measured concurrent with bird counts, it was expected that this measure would present the best opportunity for describing bird count patterns and trends.

Finally, the recently developed tree regression methodology (Chambers and Hastie 1992, Appendix D) was used to further examine the conditions under which high and low bird counts could be expected. This analysis approach combines regression models with hierarchical cluster

analysis techniques to systematically decompose the time averaged transect/cell data into groups. The groups have different overall averages and the pooled within group variance in average cell counts is much less than the total variance for all cells.

10.2 Methods

The statistical methodology used in this type of analysis is exploratory regression analysis. By regression analysis we mean the totality of the field, including model identification, model estimation and model validation. At this point in the study our concern is with model identification, examining the mathematical form of the relationship between environment and population should take. Qualitative information is examined to determine if the relationship is essentially linear or nonlinear and what are the important parameters. In the absence of a firm hypothesis as to the form of the relationship, statisticians empirically explore the use of polynomial models in all available parameters and seek to eliminate those which offer little in explaining the variability present in the population count data. Time is another factor which must be explicitly a part of the model since the goal of the model may also be to predict the state of the population at some future time.

The general structure is as follows. Let $D_{ij|t}$ be the estimated density on the j th cell of the i th transect at sample time t . Let $E_{ijk|t}$ be the value of the k th environmental parameter (index) for the j th cell of the i th transect at time t . (Note that some environmental parameters may not change over time whereas others will change greatly.)

The objective is the determination of the functional form of the relationship between these two sets of measurements. If we were to assume a functional relationship between these parameters, then

$$D_{ij|t} = f(E_{ij1|t}, E_{ij2|t}, \dots, E_{ijk|t}) + e_{ij|t}$$

Thus knowledge of the current state of the environment would be useful in predicting the expected level of population density.

It is also possible that previous states of the environment as well as population densities could be useful in describing density map changes. In this case, autocorrelations would be examined and the function could take on a time series structure.

Initial analysis examines correlations among all variables and cross-correlations over time. From this, and plots of these values, some form for the regression function, f , would be decided upon and fit using least squares or nonlinear least squares techniques.

Once a form for the model was selected and model parameters estimated, some form of validation exercise would be needed. A number of techniques are available to do this, ranging from an examination of the ability of the model to predict a part of the data which was left out in the estimation step to examination of the effectiveness of the model to predict the next time point's counts.

The final models could then be used to examine the expected change in the distribution map as a function of changes in environmental conditions. Some limited extrapolation outside the range of observed conditions could be done provided one was aware of the limitations of extrapolation from a fitted regression line.

Note that when the 'what if' or contingency analysis is performed, some thought will have to be given to the concern of a fixed overall population size being redistributed to different cells by changes in the environmental conditions, rather than allowing the population to increase or decrease. Maps of probability of occurrence or probability of cells containing a certain size class

population may be an alternative to actual predicted population density.

Other techniques could be used to classify cells as to whether they contain high or low density populations as a function of environmental parameters. This has the advantage of not depending upon the actual counts for more than to provide information on classification. In fact much of the spatial and distribution map analysis could be performed upon classification data.

10.3 Identification of land element themes

For purposes of data analysis, the total study area was divided a number of ways. A series of land elements themes were defined, each of which providing a different stratification of the study area. Land element themes are of three types: categorizations defined from existing maps (Table 2), categorizations assigned from a one-time direct observation along the sample flight path (Table 3), and multidimensional classifications created by examination of the cross-classification of all the other categorizations (Table 4). The map categorizations are for the obvious ecological parameters which might be expected to affect wading bird foraging behavior in general. The most recent maps were use, and the categorizations must be considered static for the period of study. The direct observation categorizations were also taken at one time early in the sampling program and hence they too must be considered static for the period of study. This assumption is a little more difficult to accept given the there are a number of dynamic processes (flood and fire to name just two) which can rapidly change these maps.

The multidimensional classifications are attempts to examine a more parsimonious characterization of the study area than by considering all possible cross-classifications of the land element themes. The Hierarchical Landscape Classification (HLC) is the result of a traditional geographic information system analysis of the multiple layers. The Statistical Cluster Classification (Table 6) is the result of a traditional cluster analysis of the land element themes.

10.3.1 Dominant elevations

Elevation for each cell of the 2 x 2 km gridded study area was determined by contouring known surface elevations using 1 ft. contour intervals (figure?). The 2 x 2 km cellular grid was then superimposed over the contour map. The dominate elevation for each grid cell was recorded, then assigned to a 1 ft. class elevational category. The range and spatial distribution of 1 ft elevational categories occurring in the study area are depicted in Figure 6.

10.3.2 Dominant ground surface features

Ground surface features (e.g., ponds, mud flats, creeks, etc.) were identified by direct observation along the same sample transects used for bird counts. Each of the 2 x 2 km study area grid cells were then classified to the dominate ground surface category. The categories and spatial distribution of these features occurring with the study area are presented in Figure 7.

10.3.3 Dominant soil types

Dominant soil types throughout the study area were determined by superimposing the 2 x 2 km cellular grid over U.S.D.A. Soil Conservation Service soil survey classification maps (SCS 19). Each cell was classified to its dominate soil type based on area coverage within the cell. The spatial distribution of dominate soil types is presented in Figure 8.

10.3.4 Dominant vegetation cover types

Cells were assigned to vegetative cover classes based on direct observation along the same sample aerial transects. A structural classification system (open versus dense, short versus tall, prairie versus forest, etc.) was used in assigning vegetative cover type. A dominate cover type was assigned to each 2 x 2 km grid cell based on the relative area of each type occurring in a cell.

The spatial distribution of vegetative cover is presented in Figure 9.

10.3.5 *Surface water condition*

During each survey flight, the area coverage of surface water in each grid cell was qualitatively assessed by observations made simultaneously with the bird counts. From 1985 through 1987, surface water conditions were classified into one of three categories: dry (absence of surface water), transitional, (a mixture of wet and dry) or wet (continuous surface water). From August, 1988 through the end of 1995, a five category classification scheme was used, consisting of dry/dry (absence of surface water and no groundwater visible in solution holes or ponds), dry/wet (absence of surface water but groundwater present in solution holes or ponds), dry transitional (ground surface area mostly dry but small, scattered pools of surface water present and groundwater visible in solution holes or ponds), wet transitional (ground surface area mostly wet but small, scattered dry areas present and groundwater visible in solution holes or ponds) and wet (continuous surface water over the area). For statistical analysis purposes, the 1988-95 five category metric was recoded to the 1985-87 three category metric by considering the dry/dry and dry/wet categories as dry, the dry/transitional as transitional, and the wet/transitional and wet categories as wet.

10.3.6 *Surface water hydroperiods*

Dry season hydroperiod for each study area cell was characterized for use as a land element theme in the landscape classification analysis. The number of dry season months, averaged across all study years, that each grid cell remained in the "wet" category was used as a measure of the annual dry season average hydroperiod for that cell. A robust index of each cell's long-term dry season hydroperiod characteristics was computed as the average of the dry season surface water hydroperiod values for only those years for which water levels were determined to be normal (i.e., non-extreme). This number is referred to as the trimmed mean water condition index.

10.3.7 *Landscape classification*

A hierarchical landscape classification analysis was performed following the general methodology of Astel *et al.* (1969) and implemented in GRASS 3.1 software (ref?). Land element themes were developed to identify major landscape types and more generally defined landscape categories. Land element themes for each grid cell were categorized by elevation, dominant ground surface features, substrate, trimmed mean water condition index and dominant vegetation cover type. Map layers for each land element were overlaid and categories of each combined into distinct associations, called land facets, using the 'Gcross' function in GRASS. Cells having similar land facets were then reclassified using the 'Ginfer' function to define boundaries of major landscape units and the more generally defined landscape categories. Landscape categories are defined in Table 4 and displayed in Figures 12 and 13.

10.4 *Results*

10.4.1 *Temporal Trend*

With 93 sample dates over 12 years, it is possible to describe gross annual trends in estimated abundance (computed for the whole study area), decompose this annual trend into wet/dry seasonal effects and even further explore within season trends. Annual cycles in these patterns were examined using linear regression and by computing the proportion of total annual variability which could be explained by an average annual cycle. An inter-annual wet season/dry season comparison was added to the model. The remaining within season variability was further decomposed into a within season average trend and a residual variation. Results of the variance decomposition analysis are given in Table 12.

10.4.2 *Spatial Pattern*

Within each date, bird counts were distributed spatially to the 1075 sample cells in very non-random patterns. Analysis of spatial pattern begins with an examination of general spatial aspects of bird counts, then proceeds to an exploration of the correlations of bird counts to landscape aspects and surface water condition. The measure of the importance of these factors is the fraction of total count variation explained.

Summarizing spatial pattern with one statistic is extremely difficult. In fitting the GLMM(Poisson) model to the count data, an extra dispersion parameter was estimated. This parameter is interpreted as a measure of non-randomness of counts. Large values of the extra dispersion parameter are indicative of sample dates for which guild/species counts were aggregated more than would be expected from a random spatial pattern. Similarly, small extra dispersion parameters are indicative of a more uniform distribution of counts than would be expected. Time plots of estimated overdispersion are presented in Figures 24 to 28.

10.4.3 *Correlations to Land Elements*

The degree to which land elements categorizations explain within date variability in bird counts was measured by adding each aspect as a covariate in the GLM(Poisson) model. Each land element theme examined provided clues to possible reasons for overall pattern of transect/cell counts observed. As a land element categorization was added to the model as a covariate, the unexplained within date variation was reduced. The percent reduction is presented for each land element theme in Table 13.

[Add bit about basin being the best predictor but still explaining 6.34% of residual variation (after annual cycle) at best.]

A landscape element which is not important in describing overall spatial pattern via the GLM(Poisson) model, might still provide important spatial information for specific months or even years. This possibility was examined by computing a measure of association for each landscape element theme using a Chi-square test of homogeneity statistic to examine whether relative abundance in each category followed the relative fraction of area assigned to that category. High values of this measure indicated a sampling date for which bird utilization of the landscape did not match up with the relative distribution of space defined by the landscape theme. The individual tables can be further examined to determine where bird counts were less than expected (under-utilization) and where bird counts were greater than expected (over-utilization) (see Table 14 **[Note: Table 14 was never done!]**).

10.4.4 *Correlations to Landscape Classifications*

[Summarize Table 15 here. Best explanatory power is SCC for short-legged tactile, but still only 6.31%—not much better than basin alone (6.26%).]

10.4.5 *Correlations to Surface Water Condition*

[Summarize Tables 16 and 17 here. Water alone explains 5.14% variation at best, and only 3.75% after annual cycle and basin are incorporated.]

10.4.6 *Tree Regression*

[Note: not included at this time.]

10.5 *Summary of linear modelling*

[Summarize here (paragraph or two). I can do this, based on the results in the tables, but

Ken might prefer to...

I would say something like: In the best case, the temporal model explains about 6% of total variability, and basin and water together explain about $6 + 4 = 10\%$ of the remainder, for a total of about 16%.]

11 Bird Density Prediction: A Spatial Model of Wading Bird Distribution and Abundance in Everglades National Park¹

Note: this part of the report is in preparation, and will be in the next draft (along with your comments on the remainder of *this* draft). Here's a rough abstract.

11.1 Abstract

We present a model of the foraging distribution and abundance of seven wading bird species during their nesting period (the 'dry season', December-May) in Everglades National Park (ENP). The model is based on the dynamics of surface water in the park. The data upon which the model is based are monthly aerial surveys of both birds and surface water conditions. The model is explicitly spatial, estimating the probability that a particular 2km x 2km cell will contain birds and, if it contains birds, the expected abundance. From these cell by cell estimates, we obtain predictions of the total number of occupied cells, and the total number of observed birds. Both the cell-by-cell and the overall estimates match the data very well. For some species, the model underestimates the number of birds in the driest years and overestimates the number in the wettest years *in those parts of ENP that are 'deep'* (principally the center of Shark Slough). The extremes of dryness and wetness that these dates represent are related to water management practices.

12 Overall Conclusions and Recommendations

Note: Stuart is writing this bit, and it will be in the next draft.

12.1 The future of the SRF

[Definitely!]

¹ NOTE: This Chapter is formatted as a standalone manuscript. It will provide the basis for a paper to be submitted for publication in an international peer-reviewed journal.

12.2 Possible changes in the SRF

[Do June instead of August, add November if money allows. Improve data recording efficiency (use data loggers, or palm-tops as do SFWMD).]

12.3 Looking ahead: including nesting

The loss of ephemeral wetlands that once provided late wet/early dry season shallow water foraging areas (during a time when regional water levels are highest) creates a resource bottleneck for breeding wading birds by forcing a delay in the time of their nesting, until efficient (shallow water) foraging areas are available through drying of the landscape. Insufficient time remains, however, during the dry season for nesting pairs to complete the rearing of their young, before the onset of spring rains and subsequent rising water levels that soon disperse aquatic prey. Such conditions frequently lead to nest abandonment by attending adults.

13 Acknowledgments

[Please list all names you think should be added to these!]

Financial support for this study was provided by the U.S. National Park Service, Department of Interior. David Buker and Frank Draughn, Everglades National Park, assisted in data management, mapping and analyses. Thanks are extended to these individuals for their efforts.

To collect, record, check, maintain, and analyze 12 years of data for such as large regions obviously takes a large number of researchers. At the University of Florida, Christine Steible and Frank Burnett supported the statistical analysis, Joe Aufmuth the GIS components and Patricia Cramer document review.

In Everglades National Park...

With the South Florida Water Management District...

With the Biological Research Division of USGS...

With US Fish and Wildlife Service...

14 References

[Any hints on the missing ones would be helpful, especially those related to statistics.]

Agresti, A. 1984. *Analysis of Ordinal Categorical Data*. John Wiley & Sons, Inc. p 177.

Astel *et al.* 1969.

Bancroft, G. T. 1989. Status and conservation of wading birds in the Everglades. *American Birds* 43: 1258-1265.

Bancroft, G. T., S. Jewel and A. Strong. 1991. Foraging and Nesting Ecology of Herons in the Lower Everglades in Relation to Water Conditions, Final Report to South Florida Water Management District, West Palm Beach, FL. 156 pp.

Brown. 1988.

Burhnam, K. P., Anderson, D. R. and Laake, J. L. 1976. Estimation of density from line transect sampling of biological populations. *Wildlife Monographs* 72: April. (Supplement to the *Journal of Wildlife Management* 44(2).)

DRAFT COPY

Carter et al. 1973.

Caughley, G. 1974. Bias in Aerial Survey. *Journal of Wildlife Management* 38(4): 921–933.

Caughley, G., Sinclair, R., Scott-Kemmis, D. 1976. Experiments in Aerial Survey. *Journal of Wildlife Management* 40(2): 290–300.

Chambers, J. M and T. J. Hastie. 1992. Tree-Based Models. Chapter 9 in *Statistical Models in S.* Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.

Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Measurement*. 20: 37–46.

Craighead. 1968.

Duever, M. J., J. F. Meeder, L.C. Meeder, and J. M. McCollom. 1994. The climate of south Florida and its role in shaping the Everglades ecosystem. Pages 225–248 in *Everglades: The Ecosystem and Its Restoration* (S. M. Davis and J. C. Ogden, eds.). St. Lucie Press, Delray Beach, FL.

Fleming et al. 1994.**Fleming et al. 1995.**

Frederick, P. C. and M. W. Collopy. 1988. *Reproductive Ecology of Wading Birds in Relation to Water Conditions in the Florida Everglades*. Tech. Rep. No. 30., Florida Cooperative Fish and Wildlife Research Unit, School of Forest Research and Conservation, University of Florida, Gainesville, FL. 259 pp.

GRASS. 19xx.

Gunderson and Snyder. 1994. Fire Patterns in the Southern Everglades. Pages 291–306 in *Everglades: The Ecosystem and Its Restoration* (S. M. Davis and J. C. Ogden, eds.). St. Lucie Press, Delray Beach, FL.

Harris, L. D., et al. 1994.**Johnson, R., and Fennema. 1989.**

Johnson, R. and D. Wichern. 1992. Clustering. Chapter 12 in *Applied Multivariate Statistical Analysis, Third Edition*. Prentice Hall, Englewood Cliffs, NJ.

Johnson, et al. 1993.

Jolly, G. M. 1969. Sampling methods for aerial censuses of wildlife populations. *East African Agric. Forest. J.* 34: 46–49.

Krebs, C. J. 1989. *Ecological Methodology*. Harper & Row, New York. pp654+xii.

Kushlan, J. A. and D. A. White. 1977. Nesting wading bird populations in southern Florida. *Fla. Sci.*, 40: 65–72.

Kushlan, J. A., J. C. Ogden and J. L. Tilmant. 1975. *Relation of Water Level and Fish Availability to Wood Stork Reproduction in the Southern Everglades, Florida*. Open File Report 75-434, U.S. Geological Survey, Tallahassee, FL.

Kushlan J. A., P. C. Frohring and D. Voorhees. 1984. *History and Status of Wading Birds in Everglades National Park, National Park*. Service Report, South Florida Research Center, Everglades National Park, Homestead, FL.

Leach, et al. 1972.

Light, S. S. and J. W. Dineen, 1994. Water Control in the Everglades: A Historical Perspective.

DRAFT COPY

- Pages 47–84 in *Everglades: The Ecosystem and Its Restoration* (S. M. Davis and J. C. Ogden, eds.). St. Lucie Press, Delray Beach, FL.
- Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. *SAS System for Mixed Models*. SAS Institute Inc., Cary, NC. 633 pp.
- Loftus and Ekland. 1994.**
- Ludwig, J. A. and J. F. Reynolds. 1988 [or 1992?]. *Statistical Ecology*. John Wiley & Sons, Inc. New York. pp 337+xviii.
- Macullah, P. and Nelder, J. A. 1983.** *Generalized Linear Models*. Monographs on Statistics and Applied Probability, Chapman and Hall, New York.
- McCullagh, P. and J. A. Nelder. 1989.** *Generalized Linear Models* (2nd ed.). Chapman and Hall, London.
- Nelder, J. A. and Wedderburn, R. W. M. 1972. Generalized linear models. *Journal of the Royal Statistical Society A* 135: 370–384.
- Norton-Griffiths, M., 1975. *Counting Animals*. Publication No. 1, Serengeti Ecological Monitoring Programme, African Wildlife Leadership Foundation, Nairobi, Kenya. 105 pp.
- Norton-Griffiths, M. 1978. *Counting Animals* (2nd ed.). African Wildlife Leadership Foundation, Nairobi. 110 pp.
- Noss. 1987.**
- NPS. 1990.**
- Ogden, J. C. 1978. Recent population trends of colonial wading birds on the Atlantic and Gulf coastal plains. *Proc. 1977 Conf. Colonial. Waterbird Group* 1: 155–162.
- Ogden, J. C., H. W. Kale II and S. A. Nesbitt. 1980. The influence of annual variation in rainfall and water levels on nesting by Florida populations of wading birds. *Trans. Linn. Soc. NY.* 9: 115–126.
- Parker, G. G. 1974. Hydrology of the pre-drainage system of the Everglades in southern Florida. Pages 28–27 in *Environments of South Florida: Present and Past, Memoir No. 2* (P. J. Gleason, ed.). Miami Geological Society, Coral Gables, FL..
- Pennycuick, C. J., Sale, J. B., and Stanley Price, M. 1977. Aerial Systematic Sampling applied to censuses of large mammal populations in Kenya. *E. Afr. Wildlife J.* 15: 139–146.
- Pennycuick, C. J. and Western, D. 1972. An investigation of some sources of bias in aerial transect sampling of large mammal populations. *E. Afr. Wildlife J.* 10: 175–191.
- Perkins, et al. 1993.**
- Perkins, et al. 1994.**
- Pielou, E. C. 1977. *Mathematical Ecology*. John Wiley & Sons, New York.
- Pielou, E. C. 1984. *The Interpretation of Ecological Data: A Primer on Classification and Ordination*. John Wiley & Sons, Inc. p57.
- Powell, G. V. N. 1987. Habitat use by wading birds in a subtropical estuary: Implications of hydrography. *Auk* 104: 740–749.
- Powell, G. V. N. et al. 1989.**
- Robertson, W. B. and J. A. Kushlan. 1974. The south Florida avifauna. Pages 414–452 in

- Environments of South Florida: Present and Past*, Memoir No. 2 (P. J. Gleason, ed.). Miami Geological Society, Coral Gables, FL.
- SAS Institute Inc. 1966a. The GENMOD Procedure. Chapter 10 in *SAS/STAT Software Changes and Enhancements through Release 6.11*. SAS Institute Inc., Cary, NC. 1104 pp.
- SAS Institute Inc. 1966b. The MIXED Procedure. Chapter 18 in *SAS/STAT Software Changes and Enhancements through Release 6.11*. SAS Institute Inc., Cary, NC. 1104 pp.
- SAS Institute Inc. 1982. *SAS User's Guide: Statistics* (1982 Edition). SAS Institute Inc., Cary, NC. 584pp.
- Seber, G. A. F. 1973. *The Estimation of Animal Abundance*. Hafner Publ. Co., Inc., New York, NY. pp 6-7.
- Smith, T. J. III, J. H. Hudson, M. B. Robblee, G. V. N. Powell and P. J. Isdale. 1989. Freshwater flow from the Everglades to Florida Bay: A historical reconstruction based on fluorescent banding in the coral *Solenastrea bournouni*. *Bull. Mar. Sci.* 44: 274-282.
- Snyder, G. H. and J. M. Davidson. 1994. Everglades Agriculture: Past, Present, and Future. Pages 85-116 *Everglades: The Ecosystem and Its Restoration* (S. M. Davis and J. C. Ogden, eds.). St. Lucie Press, Delray Beach, FL.
- Van Lent. 1993.
- Wanless, H. R., R. W. Parkinson and L. P. Tedesco. 1994. Sea Level Control on Stability of Everglades Wetlands. Pages 199-224 in *Everglades: The Ecosystem and Its Restoration* (S. M. Davis and J. C. Ogden, eds.). St. Lucie Press, Delray Beach, FL.
- Wolfinger, R. and M. O'Connell. 1993. Generalized linear mixed models: A pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*. 48. pp. 233-243.

15 Appendix A: Linear Model Methodology

15.1 Generalized Linear Model

A generalized linear model (a GLM) (McCullagh and Nelder 1989) was used to examine the proportion of variability attributed to temporal and spatial effects for each species. Let n_{tcym} denote the number of birds observed in the c th cell of the t th transect during the y th year and m th month. The index c runs from 1 to the number of cells in transect t , call this C_t . The index m depends on the year. Let M_y denote the set of months sampled in year y , and let L_y denote the number of months sampled in year y . The n_{tcym} are assumed to be independent random variables with an exponential dispersion family probability density (mass) function described as:

$$f(n_{tcym}; \theta_{tcym}, \phi_{tcym}) = \exp \left(\frac{n_{tcym} \theta_{tcym} - b(\theta_{tcym})}{a(\phi_{tcym})} - c(n_{tcym}, \phi_{tcym}) \right).$$

Letting

$$\theta_{tcym} = \mu_{tcym}, \quad b(\theta_{tcym}) = \frac{\mu_{tcym}^2}{2}, \quad a(\phi_{tcym}) = \sigma_{tcym}^2, \quad \text{and} \quad c(n_{tcym}, \phi_{tcym}) = -\frac{\left(\frac{n^2}{\sigma^2} + \log(2\pi\sigma_{tcym}^2) \right)}{2},$$

the probability mass function for Normally distributed data is obtained. Here, μ_{tcym} is the

expected value for the data in the c th cell of the t th transect during the y th year and m th month, and σ_{tcym}^2 is the variance. While the count is not usually assumed to follow the normal distribution, in some cases the square root or natural log of the (count + 1) might be reasonably assumed to be normal. In this paper, this is referred to this as the GLM(Normal) model. For this probability density function we have that

$$\text{Mean}(n_{tcym}) = \mu_{tcym}, \quad \text{Var}(n_{tcym}) = \sigma_{tcym}^2$$

Letting

$$\theta_{tcym} = \log(\mu_{tcym}), \quad b(\theta_{tcym}) = \mu_{tcym}, \quad a(\phi_{tcym}) = \phi_{tcym}, \quad \text{and} \quad c(n_{tcym}, \phi_{tcym}) = -\log(n_{tcym}!),$$

the probability mass function for the Poisson distributed counts is obtained. Here, μ_{tcym} is the expected value for the data in the c th cell of the t th transect during the y th year and m th month, and ϕ_{tcym} is a measure of variance called the over-dispersion parameter. The parameter ϕ measures how much the variability in the counts differs from the variability expected for observations from a Poisson distribution. In this paper, this is referred to as the GLM(Poisson) model. For this probability density function we have that:

$$\text{Mean}(n_{tcym}) = \mu_{tcym}, \quad \text{Var}(n_{tcym}) = \mu_{tcym} \phi_{tcym}$$

The GLM(Poisson) model as specified above is overparameterized since there are actually two parameters for each observation. For inference purposes, the mean and over-dispersion parameterization must be simplified. This is done by relating the expected mean count to temporal and spatial variables by equating the log of μ_{tcym} to a linear function of temporal and spatial regressors. That is:

$$\log(\mu_{tcym}) = \alpha + \beta_1 y + \beta_2 y^2 + \beta_3 \sin\left(\frac{2\pi}{12} m\right) + \beta_3 \cos\left(\frac{2\pi}{12} m\right) + s(t, c)$$

and

$$\phi_{tcym} = \phi \quad (\text{a constant}).$$

This model dictates that the log of the mean number of birds per cell is modeled by three components. The first component represents the overall spatial and temporal average count. The second and third components ($\beta_1 y + \beta_2 y^2$) measure long-term trends in average count modeled as a linear and quadratic function of year. The fourth and fifth terms use sine and cosine functions to model a sinusoidal cycle having a period of twelve months. These terms assess inter-year trends in monthly counts averaged over all transect and cell counts. The last component ($s(t, c)$) refers to a general spatial function defined for all values of t and c . Its purpose is to assess the impact of a spatial categorization or classification, such as basin or pond density, on the expected count. The first two components combined measure the proportion of total count variability which can be attributed to temporal effects. The last component in the model allows us to measure the proportion of total count variability attributed to spatial effects which can be explained by a spatial categorical or classification variable.

Thus, the proportion of variability in bird counts can be partitioned into temporal and spatial effects by fitting reparameterized GLM(Poisson) models of increasing complexity, by adding the above components to the model one at a time. In normal model analysis of variance, the differences between the residual deviance (error) would be called the Type 1 sums of squares. The sequence in which variables were added to the model is: intercept, year, month, and spatial variable. Table 1 shows an example of an analysis of deviance table. For this example, the proportion of variability explained by temporal effects is calculated in the following manner. First the difference in deviances between the intercept-only model and the model containing the year and month terms is computed ($100 - 70 = 30$). Second, this difference in deviances is divided by

the deviance of the intercept model ($30 / 100 = 0.3$). This final number (0.3) is the proportion of variability explained by temporal effects. The proportion of variability explained by spatial effects when temporal effects have been accounted for is calculated in a similar fashion. First, the difference in deviances between the full model (year, month, and spatial variable) and the model with only year and month terms is calculated ($70 - 50 = 20$). Next, this difference in deviances is divided by the deviance of the intercept only model ($20 / 100 = 0.2$). The final number (0.2) is the proportion of variability explained by the spatial variable when temporal effects have been accounted in the model. This can be written in a table as given below.

Variable	Deviance
Intercept	100
Year	80
Month	70
Spatial Variable	50

Computations for fitting the GLM(Poisson) models were performed using the SAS procedure GENMOD (SAS Institute Inc 1996). The GLM(Normal) model was not used in this paper, but would be fit using the SAS procedure, GLM (SAS Institute Inc 1996).

15.2 Generalized Linear Mixed Model

A generalized linear mixed model (GLMM) (Wolfinger and O'Connell 1993) is a broader class of generalized linear models. They were used in this paper to 1) determine the nature of spatial correlations to include in the abundance estimation model and 2) to estimate bird abundance for each species and date combination.

The initial screening models were developed from the GLM(Normal) model. In this case the $z_{t_{cym}} = \log_e(n_{t_{cym}} + 1)$ were assumed to follow a normal family of probability density functions with dependency in transect and cell observations taking a number of forms (see Table 5). The different models specified quite different correlation structures as follows:

Model Name	Correlation Structure
Independence	$\text{Var}(z_{t_{cym}}) = \sigma_{ym}^2$ $\text{Cov}(z_{t_{cym}}, z_{t'c'ym}) = 0 \quad \text{for } t = t', c = c' \quad \text{for } t = t', c = c'$
Compound Symmetry, within a transect	$\text{Var}(z_{t_{cym}}) = \sigma_{ym}^2 + \phi_{ym}^2$ $\text{Cov}(z_{t_{cym}}, z_{t'c'ym}) = \phi_{ym}^2 \quad \text{for } t = t', c \neq c'$
AR(1) within independent	$\text{Var}(z_{t_{cym}}) = \sigma_{ym}^2$

transects	$\text{Cov}(z_{tcym}, z_{t'c'ym}) = \sigma_{ym}^2 \phi_{ym}^{ c-c' }$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = 0$	for $t = t', c \neq c'$ for $t \neq t'$
AR(1) with random transects	$\text{Var}(z_{tcym}) = \sigma_{ym}^2 + \psi_{ym}^2$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = (\sigma_{ym}^2 + \psi_{ym}^2) \phi_{ym}^{ c-c' }$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = \psi_{ym}^2$	for $t = t', c \neq c'$ for $t \neq t'$
Spherical, no nugget effect	$\text{Var}(z_{tcym}) = \sigma_{ym}^2$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = 1 - 1.5 \left(\frac{d_{tct'c'}}{\rho_{ym}} \right) + 0.5 \left(\frac{d_{tct'c'}}{\rho_{ym}} \right)^2$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = \sigma_{ym}^2$ where $d_{tct'c'} = \sqrt{(t-t')^2 + (c-c')^2}$	for $d_{tct'c'} < \rho_{ym}$ for $d_{tct'c'} \geq \rho_{ym}$
Spherical, with a nugget effect	$\text{Var}(z_{tcym}) = \sigma_{ym}^2 + \psi_{ym}^2$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = 1 - 1.5 \left(\frac{d_{tct'c'}}{\rho_{ym}} \right) + 0.5 \left(\frac{d_{tct'c'}}{\rho_{ym}} \right)^2$ $\text{Cov}(z_{tcym}, z_{t'c'ym}) = \sigma_{ym}^2$ where $d_{tct'c'} = \sqrt{(t-t')^2 + (c-c')^2}$	for $d_{tct'c'} < \rho_{ym}$ for $d_{tct'c'} \geq \rho_{ym}$

For the final estimation model, the n_{tcym} were assumed to follow the same exponential family probability density function described for GLM(Poisson), but a dependency in the transect and cell observations for a particular sample date is modeled using a covariance structure given by:

$$\begin{aligned} \text{Var}(n_{tcym}) &= \sigma_{ym} + \phi_{ym} \\ \text{Cov}(n_{tcym}, n_{t'c'ym}) &= \mu_{ym} \phi_{ym} \phi_{ym}^{|c-c'|} && \text{for } t = t' \\ \text{Cov}(n_{tcym}, n_{t'c'ym}) &= 0 && \text{for } t \neq t' \end{aligned}$$

This model specifies that observations within the same transect are correlated and cell counts along a transect that are closer together are more correlated than observations farther apart. Cell counts of two different transects are assumed independent. When $\phi = 0$, this model simplifies to the GLM(Poisson). We refer to this as the GLMM(Poisson) model. The SAS macro GLIMMIX (Littell, Milliken, Stroup, and Wolfinger 1996) was used to compute estimates of the unknown parameters μ , ϕ and σ . The parameter estimation procedure also provides standard error estimates for these parameter estimates. Using the estimate for average density, μ , bird abundance, a_{ym} , for the y th year and m th month was calculated by:

$$a_{ym} = \frac{N\mu_{ym}}{p_c}$$

where $N = 1075$ is the number of cells sampled on one date and $p_c = 0.15$, representing a 15% coverage of each cell in the aerial sampling. A confidence interval for a was also computed.

16 Appendix B: Statistical Cluster Classification Development

The land element themes (Tables 2 and 3) contain different categorizations of the 1,075 cells in the study and hence each provides a separate spatial characterization. The Statistical Cluster Classification (Table 4) was developed using standard statistical cluster analysis methods applied to eight of the land element themes (elevation, soil type, specific vegetation, basin, physiographic region, watershed, pond density, and ground surface features). The goal in developing this classification was to create groups such that cells within a class have similar land element theme values, and at the same time, cells in different classes differ as much as possible in as many of these land element categories as possible.

A two-step cluster procedure was used to create the classifications. In the first step, a hierarchical clustering procedure was used to create an initial set of cluster groups. This method uses a measure of distance or similarity to group cells. Since all eight land element variables in the analysis are categorical, a similarity measure based on indicator variables for land element variable levels was used. Thus, for example, seven indicator variables were created for elevation (call them e_1 to e_7). If a cell was classified in the i th elevation category, then e_i was set to 1, otherwise $e_i = 0$. Distance between two cells was defined as Euclidean distance using all of the indicator variables. As an example, suppose the distance between two cells is based on physiographic region and pond density. The table below shows physiographic region and pond density along with the associated indicator variables for two hypothetical cells. Ph_1 – Ph_3 and Po_1 – Po_3 are the indicator variables associated with physiographic region and pond density respectively.

Cell	Physiographic Region	Pond Density	Ph_1	Ph_2	Ph_3	Po_1	Po_2	Po_3
1	1	2	1	0	0	0	1	0
2	1	3	1	0	0	0	0	0

The Euclidean distance between cells 1 and 2 is

$$\left[(1-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2 + (0-1)^2 \right]^{\frac{1}{2}} = 2^{\frac{1}{2}}$$

This distance function counts the number of land element variables for which cells differ, multiplies the count by two, and takes the square root of the product.

The hierarchical clustering technique starts with 1075 groups. Each group contains one observation. The two groups with the smallest distance are merged into one group. The merging process continues until it is deemed prudent to stop the process. Distance between groups is computed as the average distance between all pairs of cells between the two groups.

A characteristic of the hierarchical clustering implemented in this analysis is that once a cell is classified into a group, the cell can never leave the group. In the second step, the initial cluster groups were reconstructed using a 'k-means' clustering algorithm (Johnson and Wichern 1992) which produces groups which were internally more consistent. The input to the k-means procedure is the number of groups created from the hierarchical procedure and each group's centroid (vector of majority values for the category code for each land element). The k-means procedure calculates the distance between a cell and the centroid of all the groups and assigns the cell to the group with the nearest centroid.

17 Appendix C: Computing drying (P_D) and wetting (P_W) proportions and the change potential index (CPI).

Let n_{ij} be the number of cells meeting both row and column conditions. That is, n_{11} is the number of cells which were dry in both Month 1 and Month 2. These counts can be arranged in the following table.

		Month 2		
		Dry	Transitional	Wet
Month 1	Dry	n_{11}	n_{12}	n_{13}
	Transitional	n_{21}	n_{22}	n_{23}
	Wet	n_{31}	n_{32}	n_{33}

$$P_D = (n_{21} + n_{31} + n_{32}) / (n_{21} + n_{22} + n_{23} + n_{31} + n_{32} + n_{33})$$

$$P_W = (n_{12} + n_{13} + n_{23}) / (n_{11} + n_{12} + n_{13} + n_{21} + n_{22} + n_{23})$$

$$CPI = P_D - P_W$$

18 Appendix D: Tree Regression Methodology

A tree regression analysis was performed on the count data for each species/guild combination to determine if there were cells which demonstrated consistently high or low counts over time. The analysis was performed on the sum of cell counts for dry season months only. Let $n_{tc..}$ represent the total dry season count for transect t , cell c . Let

$$\mathbf{X} = \left\{ X^{(1)}, X^{(2)}, \dots, X^{(p)} \right\}_{(t \times c) \times p}$$

denote a vector of the predictor variables, in this case the land element themes, categorizations, and classifications. The measurement space, denoted Ω , for $\mathbf{X}$ is defined as the set of all possible combinations of values for the vector $\mathbf{X}$. In the analyses using the GLM(Poisson) and GLMM(Poisson) models, the objective was to examine the relationship between n_{tcym} and $\mathbf{X}$ using a real-valued predictor function defined on Ω .

When the exact form of the predictor equation is not known, and the dimensionality of X is such that some form of variable selection is required. Assuming a linear predictor, either a stepwise or best subsets selection algorithm is usually recommended. In the case of the GLMM(Poisson) model, this process did not result in increases in the percent variability explained. In addition, the complexity of the relationships assumed increases dramatically once one increases the number of covariates in the model.

Tree-based regression was used as an alternative to the normal variable selection methodologies. Like other robust regression techniques, tree-based regression facilitates development of predictors when its general form is not known.

In tree-based regression, the measurement space, Ω is partitioned by repeated binary splitting of subsets of Ω into two descendent subsets, starting with Ω itself. The partition of Ω is usually represented as a binary tree (Figure xxx), denoted T , with nodes, T_k , $k = 1, 2, \dots, K$. This binary recursive partitioning of Ω is done in such a way that the within-node variability in observed response (total cell count) is minimized. That is, the added residual variance of the two descendent nodes is less than the residual variance in the original ancestor.

If we let Ω_k be that subset of Ω represented in the k th node, then the value of the predictor function for node k is the mean of the observed responses in that node. If one designates the set of all transect/cells in the node T_k by $\{(t,c) \in T_k\}$, then the average response for this node is

$$f(x|x \in \Omega_k) = \bar{n}(x) = \bar{n}(T_k) = \frac{\sum_{(t,c) \in T_k} n_{tc..}}{\sum_{(t,c) \in T_k} 1}$$

In this way, the prediction surface (i.e., expected total cell counts) is approximated by a step function defined on all transect/cell combinations, whose value depends on which node, T_k the combination falls into. Each subset of Ω_k really translates to a set of transect/cells, each having similar values for some subset of land element themes. The model assigns to each of these transect/cells a expected count value which is the average of all total counts all included transect/cells.

The residual sums of squares (or deviance) associated with $f(x)$, for the sample (n_{tc}, x_{tc}) , is

$$R(f) = \sum_{(t,c)} (n_{tc..} - f(x_{tc}))^2$$

For node T_j of tree T , define the associated residual sums of squares or residual deviance as:

$$r(T_j) = \sum_{(t,c) \in T_j} (n_{tc..} - \bar{n}(t_j))^2$$

The residual sums of squares for tree T is the sum of all terminal node-associated sums of squares. That is, if T_j , $j = 1, 2, \dots, k^*$ are the terminal nodes of tree T , then

$$R(T) = \sum_{j=1}^{k^*} \sum_{(t,c) \in T_j} (n_{tc..} - \bar{n}(t_j))^2$$

The measure of improvement observed by splitting node T_j into nodes $T_{j'}$ and $T_{j''}$ is given by

$$r(T_j) - r(T_{j'}) - r(T_{j''}).$$

Using this measure of improvement, the best split at a node is that split which most successfully separates the high response values from the low ones. That is, for a successful split,

$$\bar{n}(T_{j'}) < \bar{n}(T_j) < \bar{n}(T_{j''}) \text{ and } R(\text{unsplit tree}) > R(\text{split tree})$$

Candidates for splits are defined through the individual components of X . Let X be composed of p components, that is the t th observation is

$$x'_{tc} = \{x^{(1)}_{tc}, x^{(2)}_{tc}, \dots, x^{(p)}_{tc}\}.$$

For an ordinal, interval or ratio component, say $X^{(l)}$, splits can be created as

$$T_{j'} = \{(t, c) | x_{tc}^{(l)} < x^0\}$$

$$T_{j''} = \{(t, c) | x_{tc}^{(l)} > x^0\}$$

for each observed value of x^0 in the domain of $X^{(l)}$. Similarly, if $X^{(l)}$ is a nominal scale variable with observed levels $C_l = \{C_{l1}, C_{l2}, \dots, C_{lL}\}$, then splits are created as

$$\begin{aligned} T_{j'} &= \{(t, c) | x_{tc}^{(l)} \in C', C' \subset C\} \\ T_{j''} &= \{(t, c) | x_{tc}^{(l)} \in \bar{C}', \bar{C}' \subset C, \bar{C}' \cap C = \emptyset, \bar{C}' \cup C' = C\} \end{aligned} \quad (6)$$

For each subset Ω_k , the split which maximizes the improvement in residual sums of squares is chosen from among all possible splits across all components. In this way, each component is examined at each opportunity to split. This partitioning of Ω is usually performed until either 1) each terminal node has fewer than a prespecified number of observations (usually 5), or 2) the variability in each terminal node is less than some small value, usually defined to be 1% of the variance of the root node. This results in a tree which is much larger than the data can support. As with over-specification in multiple linear regression, the final predictor function is over-optimistic in that the estimate of residual variance is much smaller than it should be.

To correct for over-optimistic trees, the maximum-sized tree is pruned back to a smaller tree. Our objective is to simplify the tree without significantly sacrificing goodness-of-fit. The pruning process produces a nested sequence of trees determined by recursively snipping off the least important branches. Importance is measured by a cost-complexity index

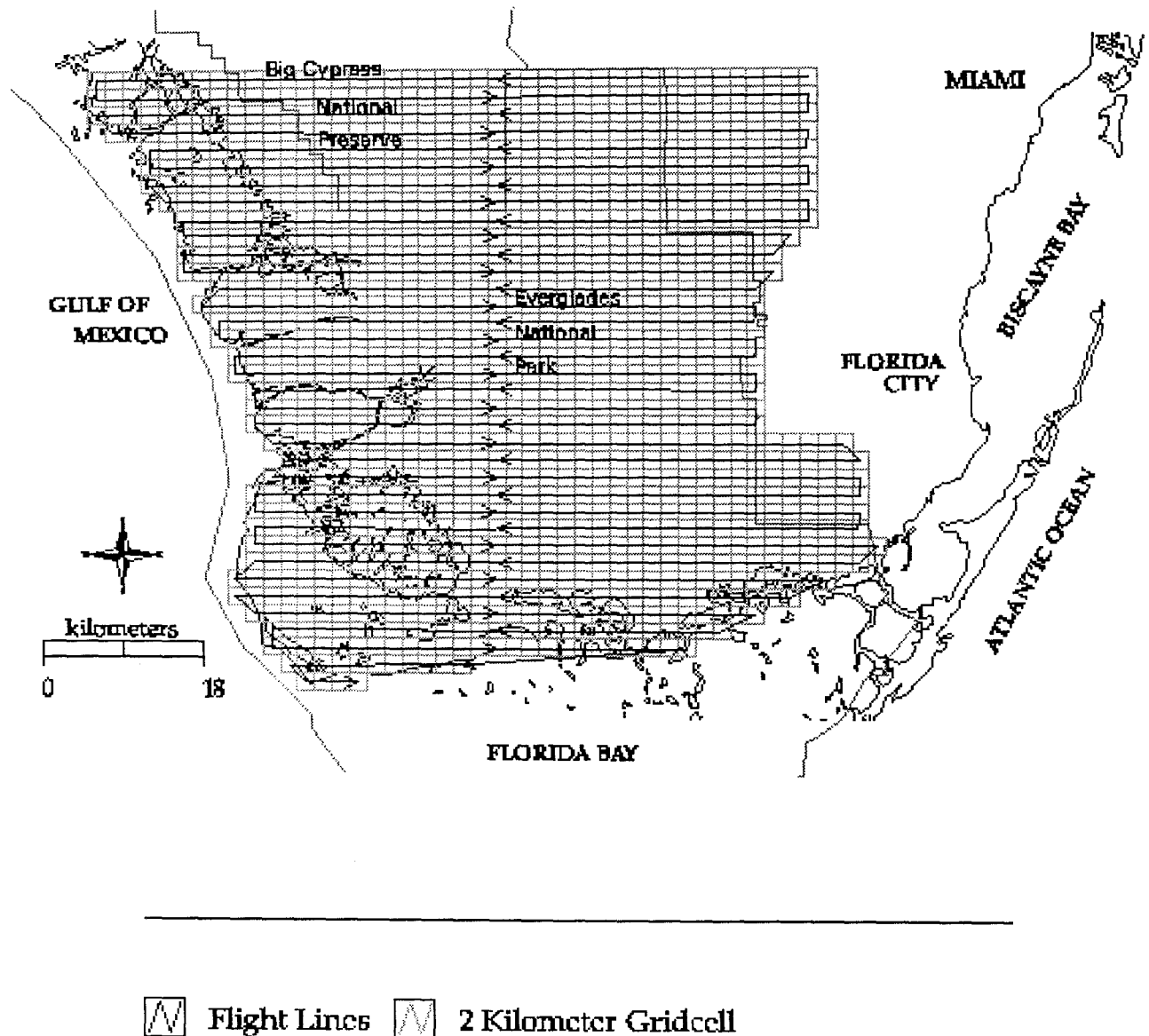
$$R_\alpha = R(T) + \alpha|T|$$

where $|T|$ = number of nodes in tree T , and α is the cost-complexity parameter. For a given value of α , the optimal subtree minimizes the cost-complexity index. Optimal subtree residual variance decreases and tree size increases monotonically as α decreases. Venables and Ripley (1994) use Mallows's C_p statistic to suggest using $\alpha = K\hat{\sigma}^2$ where $\hat{\sigma}^2$ is the residual mean square for the maximum size tree and K is a value between 2 and 6.

Supplemental information on optimal tree size is obtained by examination of tree prediction error using a 10-fold cross-validation process. Here the data are divided into 10 roughly equal parts and a tree is developed using 9 parts with prediction error (residual variances) computed on the tenth part. This is done in the 10 possible ways, with prediction error averaged for optimal subtrees for specified values of α . A plot of prediction error by optimal subtree size is not necessarily monotonically decreasing but trees with minimum prediction error are candidates for optimal tree.

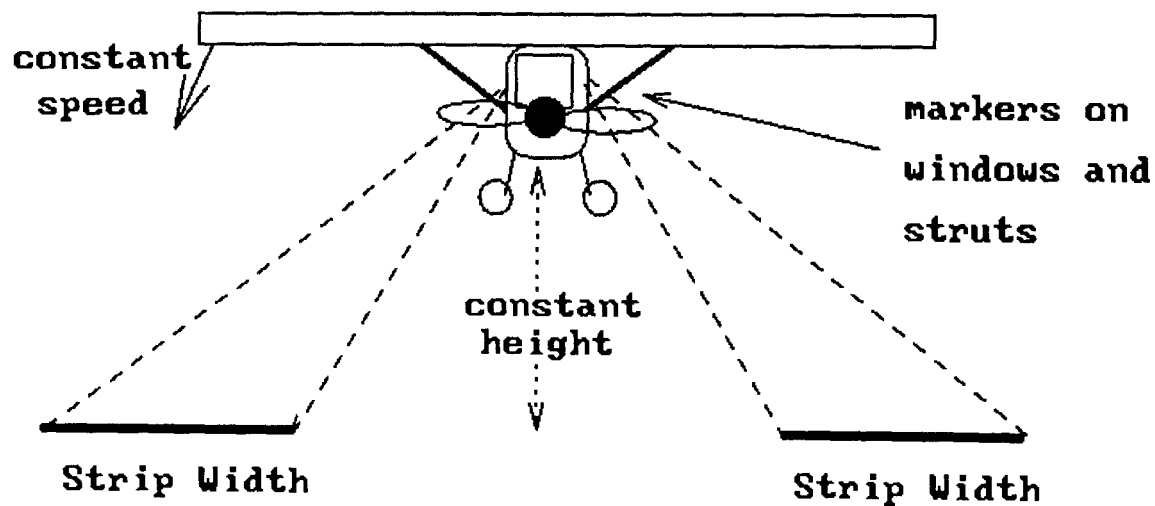
17 Figures

17.1 Figure 1. Map of southern Everglades and Big Cypress National Preserve with sampling grid and flight line



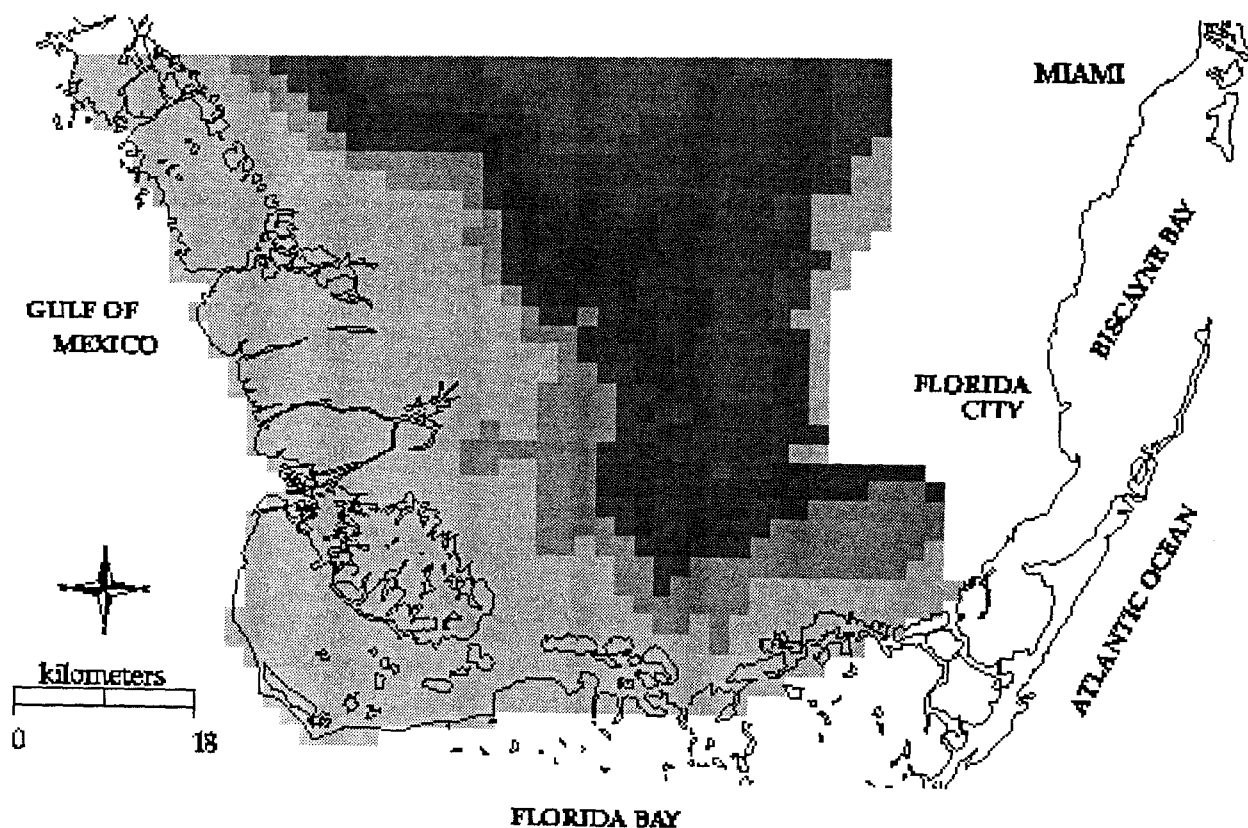
DRAFT COPY

17.2 Figure 2. SRF sampling protocol. (Figure of fixed wing aircraft with strip width, height, speed)



DRAFT COPY

17.3 Figure 3. Dominant elevations in the SRF study area

**Observed Class:**

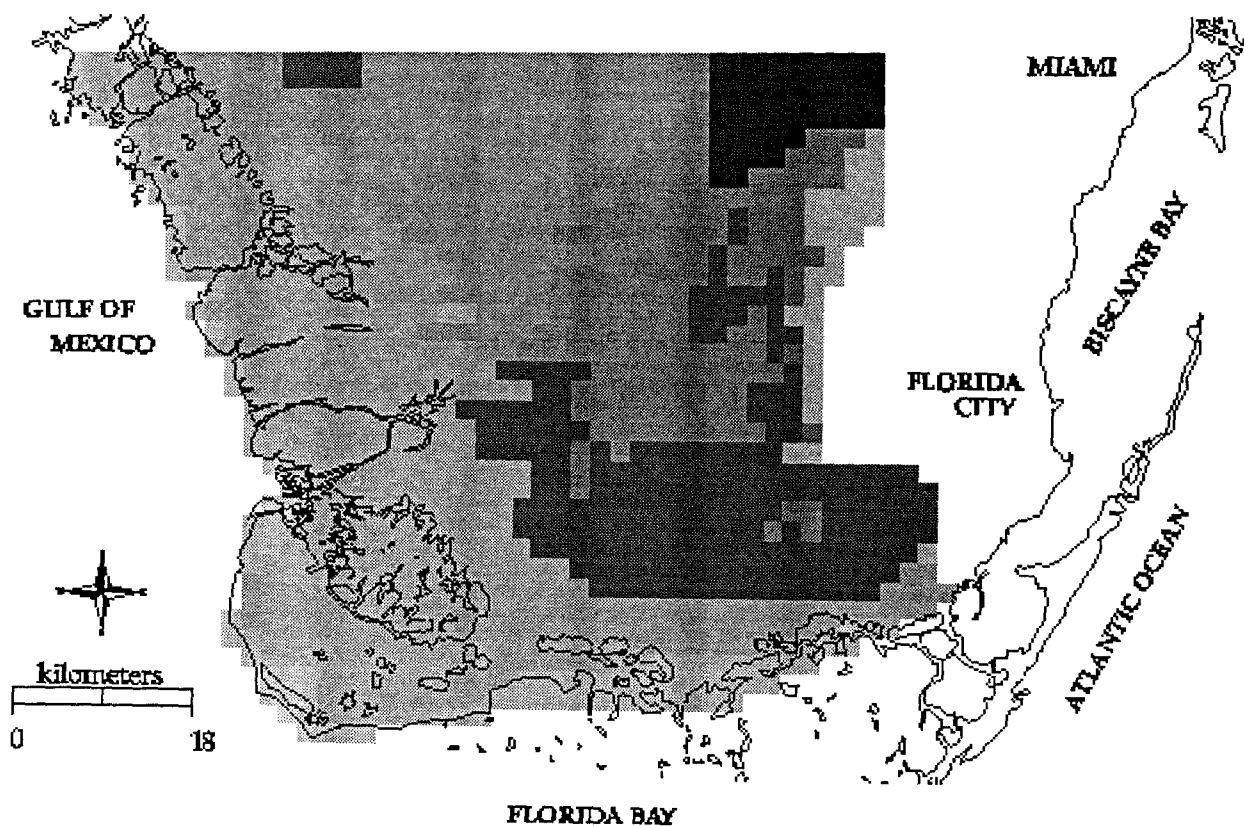
■ Incomplete Temporal Record,
Not In Estimate

Elevation (ft. msl)

■	$0 < X \leq 1$
■	$1 < X \leq 2$
■	$2 < X \leq 3$
■	$3 < X \leq 4$
■	$4 < X \leq 5$
■	$5 < X \leq 6$
■	$6 < X \leq 7$

DRAFT COPY

17.4 Figure 4. Dominant soil types in the SRF study area



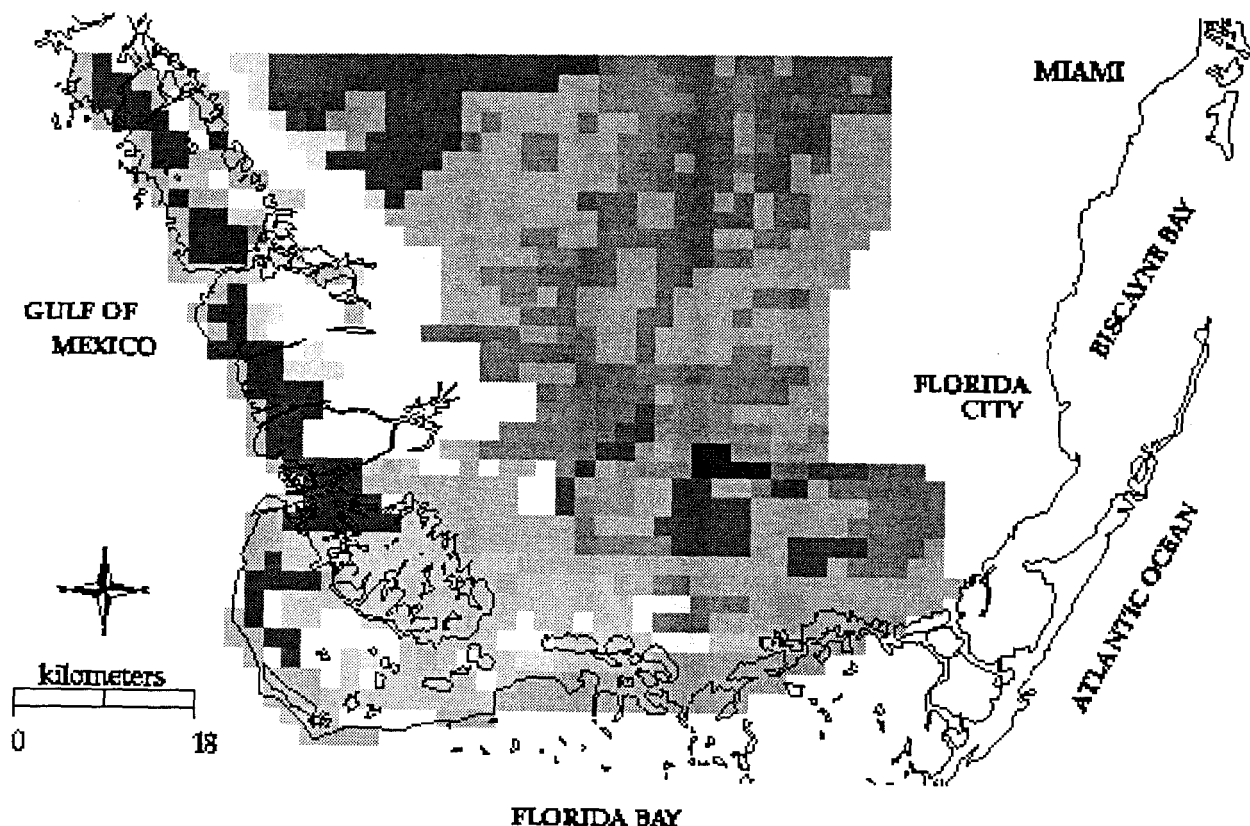
Observed Class: ☐ Incomplete Temporal Record, Not In Estimate

Soil Types

- | | |
|---|---------------------------------------|
| ☐ Marls and/or Mucks, Tidal (unclassified) | ☐ Marls, Sandy |
| ☐ Mucks, Freshwater | ☐ Sands |
| ☐ Marls/Mucks, Freshwater (mixed) | ☐ Rockland |
| ☐ Marls, Organic and/or Calcium | |

DRAFT COPY

17.5 Figure 5. Dominant vegetation in the SRF study area



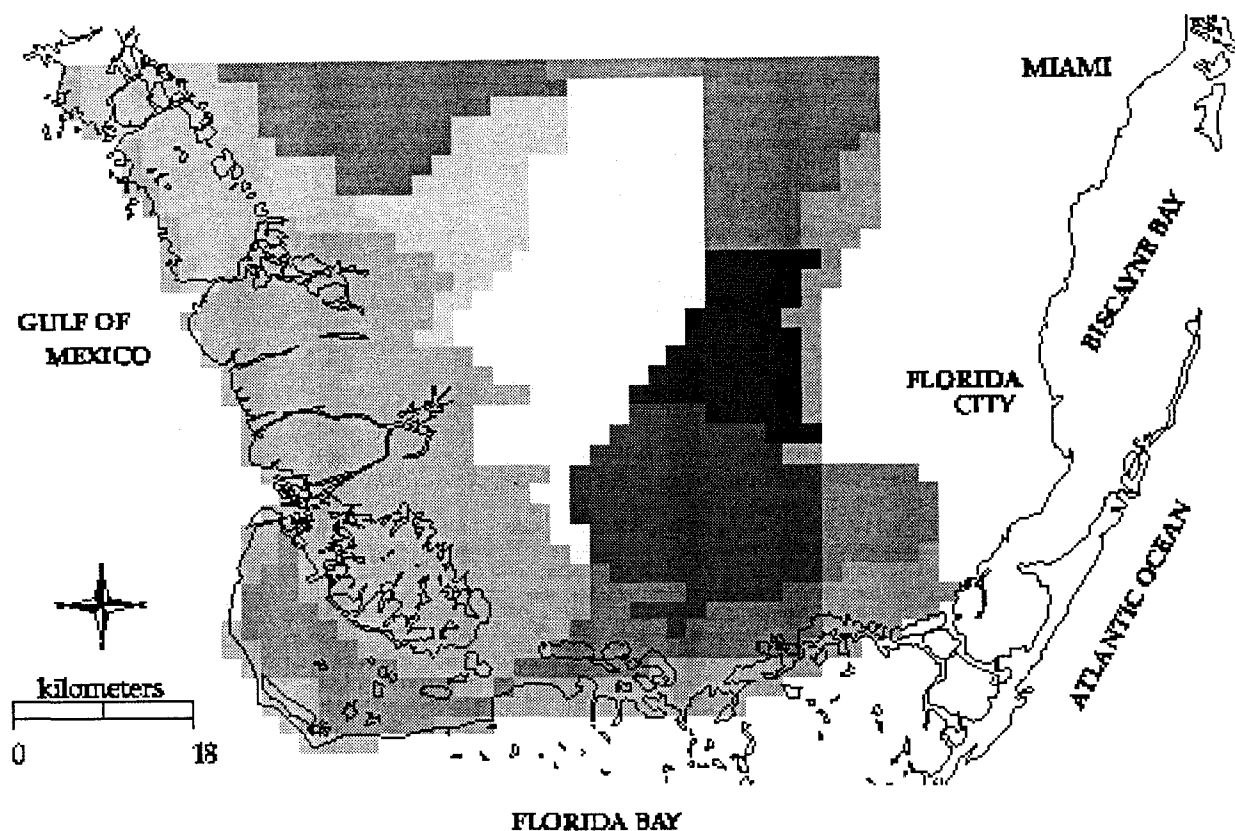
Observed Class: ☐ Incomplete Temporal Record, Not In Estimate

Vegetation Cover Type

- | | |
|---|---|
| ☐ Tall Mangrove Forest | ☐ Open Sawgrass Prairie / Muhlenbergia Prairie |
| ☐ Tall/Short (Mixed) Mangrove Forest | ☐ Tree Island/Open Prairie |
| ☐ Short Mangrove Forest | ☐ Swamp Forest |
| ☐ Mangrove/Prairie (Mixed) | ☐ Cypress Forest |
| ☐ Coastal Prairie | ☐ Cypress/Pine (Mixed) Forest |
| ☐ Open Water/Slough | ☐ Pineland |
| ☐ Dense Sawgrass Prairie | ☐ Agricultural/Exotics |
| ☐ Open Sawgrass Prairie | |

DRAFT COPY

17.6 Figure 6. Basin regions for the SRF study area



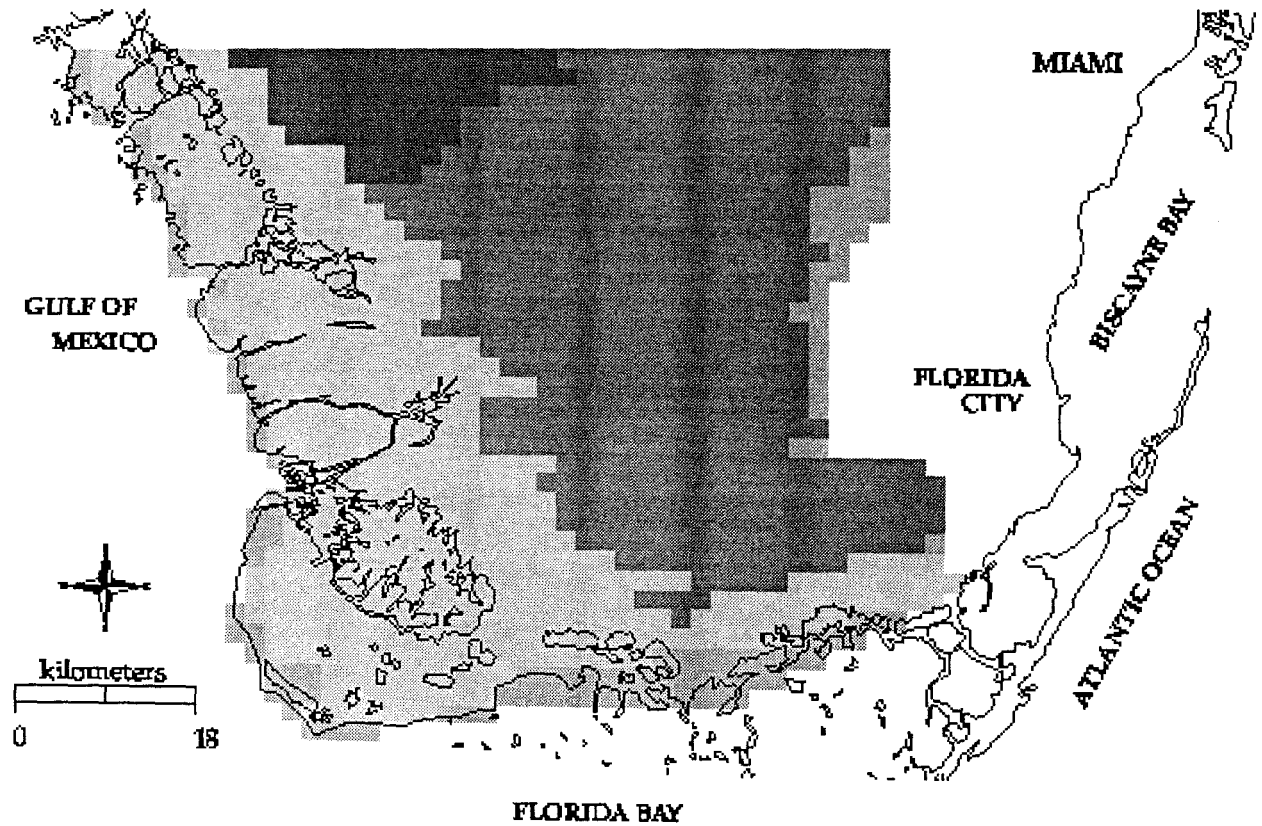
Observed Class: ☐ Incomplete Temporal Record, Not In Estimate

Drainage Basin

- | | |
|---|--|
| ☐ WCA 3-A South | ☐ Shark Slough Mang. Est. |
| ☐ Northeast Shark Slough | ☐ Shark Slough |
| ☐ Northern Taylor Slough | ☐ East Slough |
| ☐ LPK/South Taylor/ Geo. Area | ☐ Big Cypress Mang. Est. |
| ☐ Eastern Panhandle | ☐ Southern Big Cypress |
| ☐ Eastern Panhandle Mang. Est. | |
| ☐ LPK/STS/EP/ Mang. Est. | |
| ☐ Cape Sable | |

DRAFT COPY

17.7 Figure 7. Physiographic regions for the SRF study area



Observed Class:

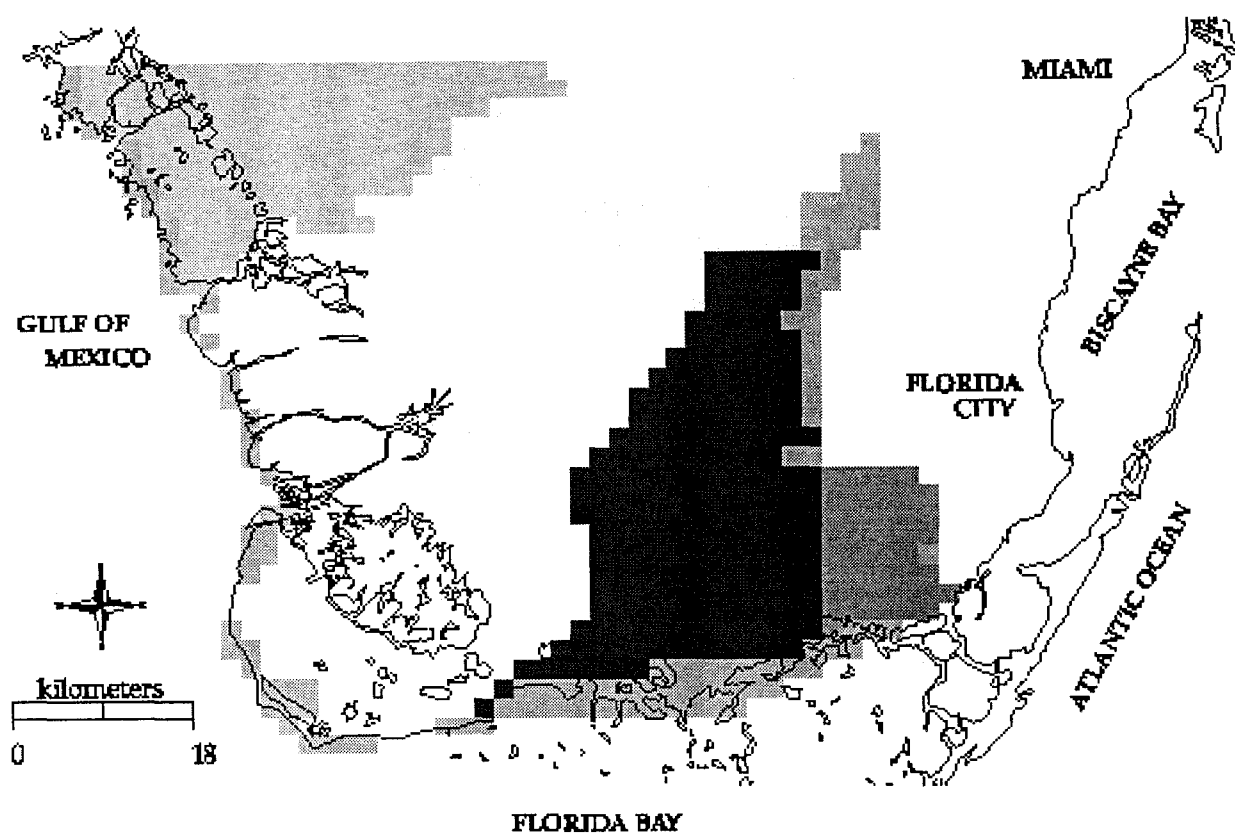
- Incomplete Temporal Record,
Not In Estimate

Physiographic Regions

- Mangrove Estuaries
- Freshwater Grassland
- Cypress Forest

DRAFT COPY

17.8 Figure 8. Watershed regions for the SRF study area

**Observed Class:**

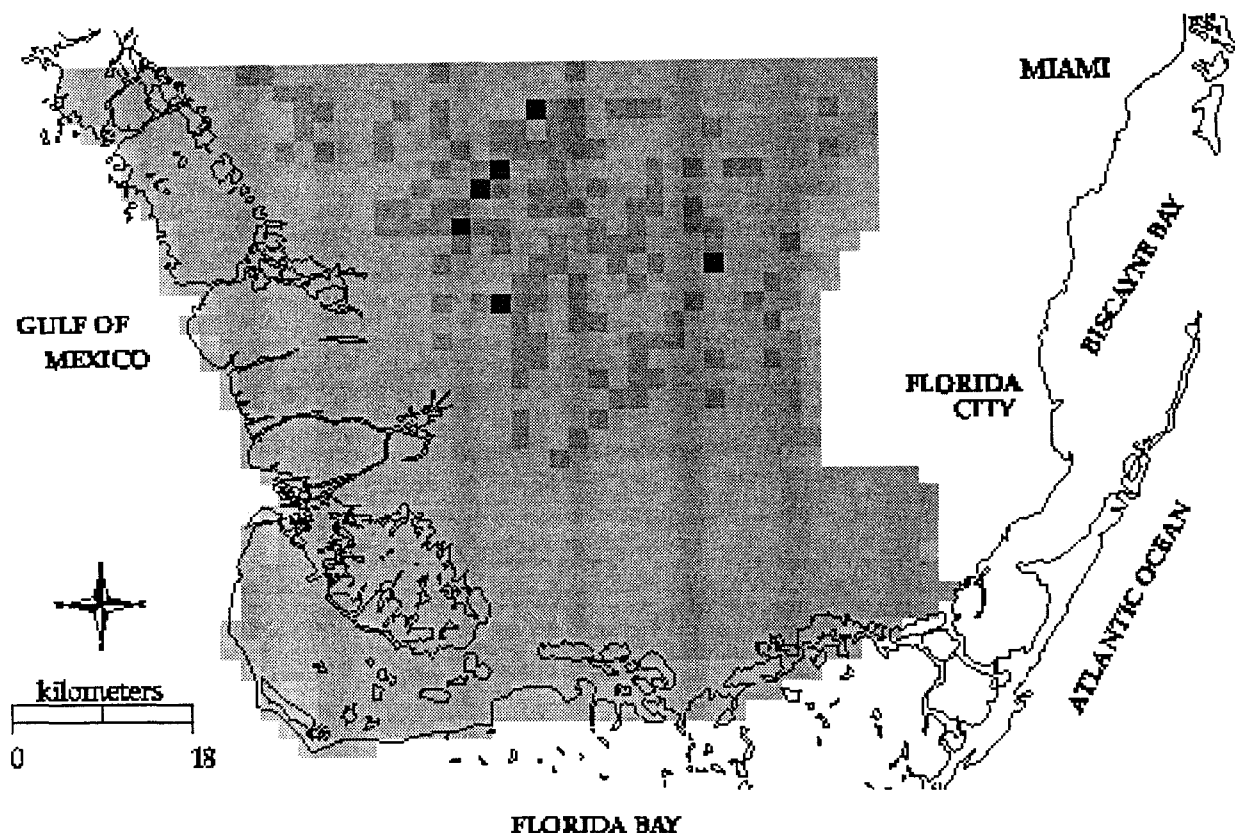
- Incomplete Temporal Record,
Not In Estimate

Watersheds

- Big Cypress Watershed
■ Eastern Panhandle Watershed
□ Shark Slough Watershed
■ Taylor Slough Watershed

DRAFT COPY

17.9 Figure 9. Pond density classification for the SRF study area



Observed Class:

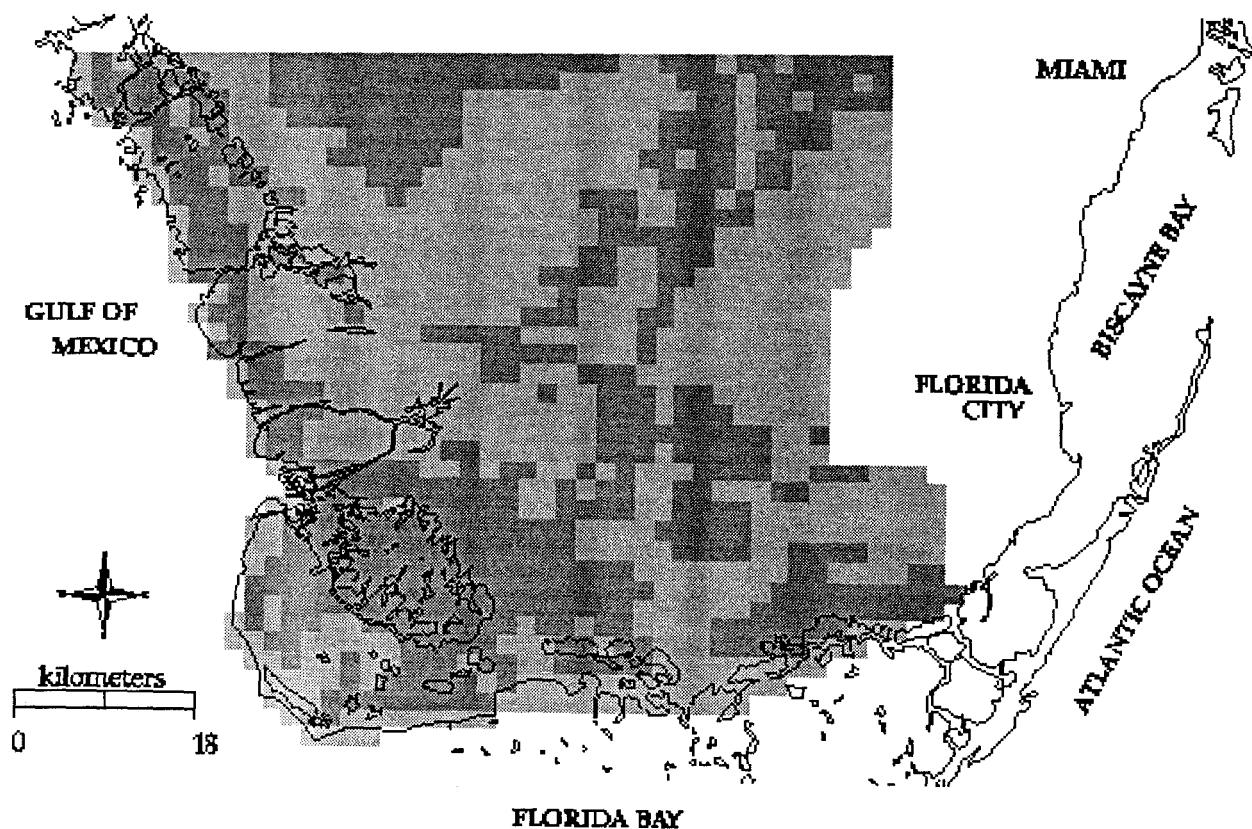
■ Incomplete Temporal Record,
Not In Estimate

Pond Density

■ No Ponds
■ Low Density
(1-5/km²)
■ Moderate Density
(6-10/km²)

DRAFT COPY

17.10 Figure 10. Gross vegetation classification for the SRF study area

**Observed Class:**

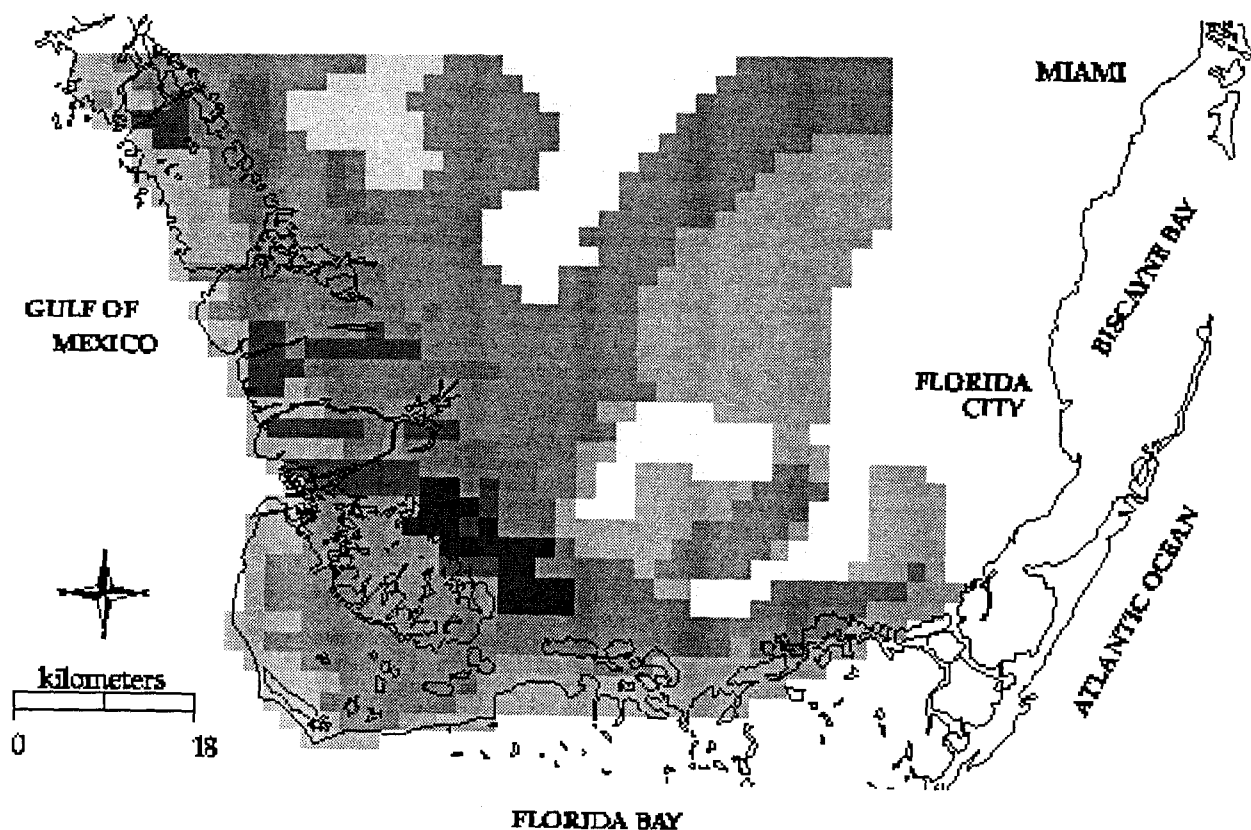
- Incomplete Temporal Record,
Not In Estimate

Vegetation Densities

- Dense Vegetation
■ Open Vegetation

DRAFT COPY

17.11 Figure 11. Dominant ground surface features for the SRF study area



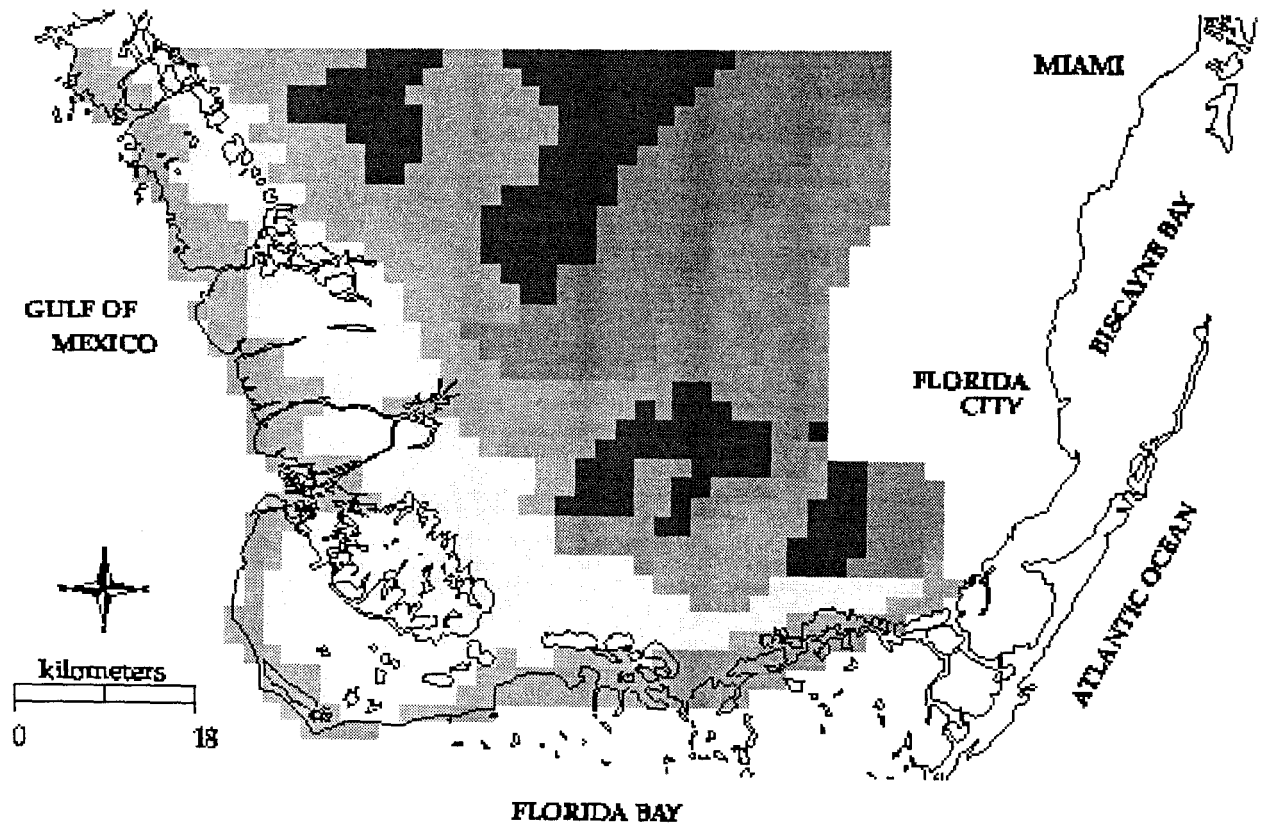
Observed Class: ☐ Incomplete Temporal Record, Not In Estimate

Ground Surfaces

- | | | |
|---------------------------------------|--|----------------------------------|
| ☐ Beachrim | ☐ Tidal Creeks | ☐ Lowland |
| ☐ Deep Bay | ☐ Coastal Tidal Flats | ☐ Ridge |
| ☐ Shallow Bay | ☐ Inland Flats | ☐ Outcrop |
| ☐ Tidal Rivers | ☐ Slough | |

DRAFT COPY

17.12 Figure 12. Hierarchical Landscape Classification of the SRF study area



Observed Class:

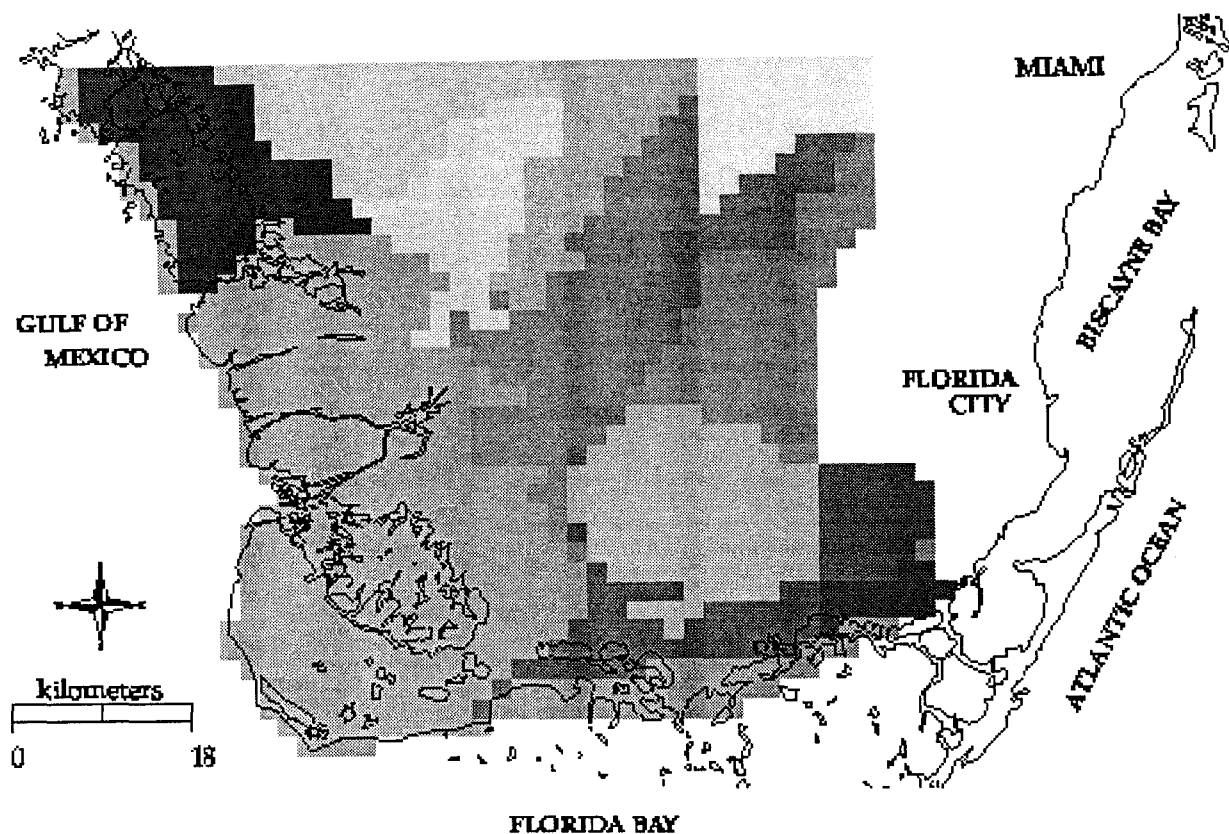
- Incomplete Temporal Record,
Not In Estimate

Landscape Categories

- Outcrop/Ridge Complex
- Rockland
- High Slough/Swamp Complex
- Central Sloughs
- Estuary Headwaters
- Estuary Interior Flats/Bay Complex
- Beach Rim/Tidal Flat Complex

DRAFT COPY

17.13 Figure 13. Statistical Cluster Classification of the SRF study site

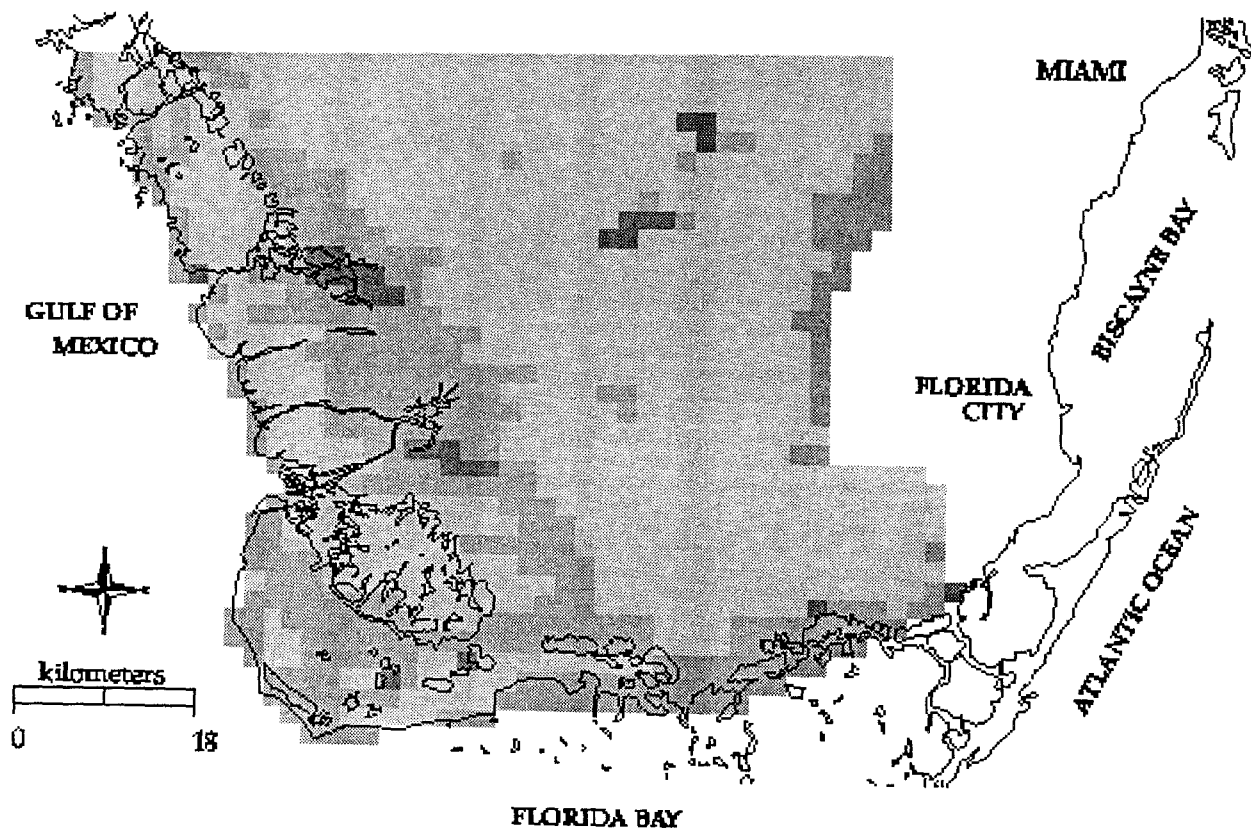


Observed Class: Incomplete Temporal Record, Not In Estimate

Statistical Cluster Class

 Class 1	 Class 6	 Class 11
 Class 2	 Class 7	 Class 12
 Class 3	 Class 8	 Class 13
 Class 4	 Class 9	 Class 14
 Class 5	 Class 10	

17.14 Figure 14. Example dry month and wet month water condition maps

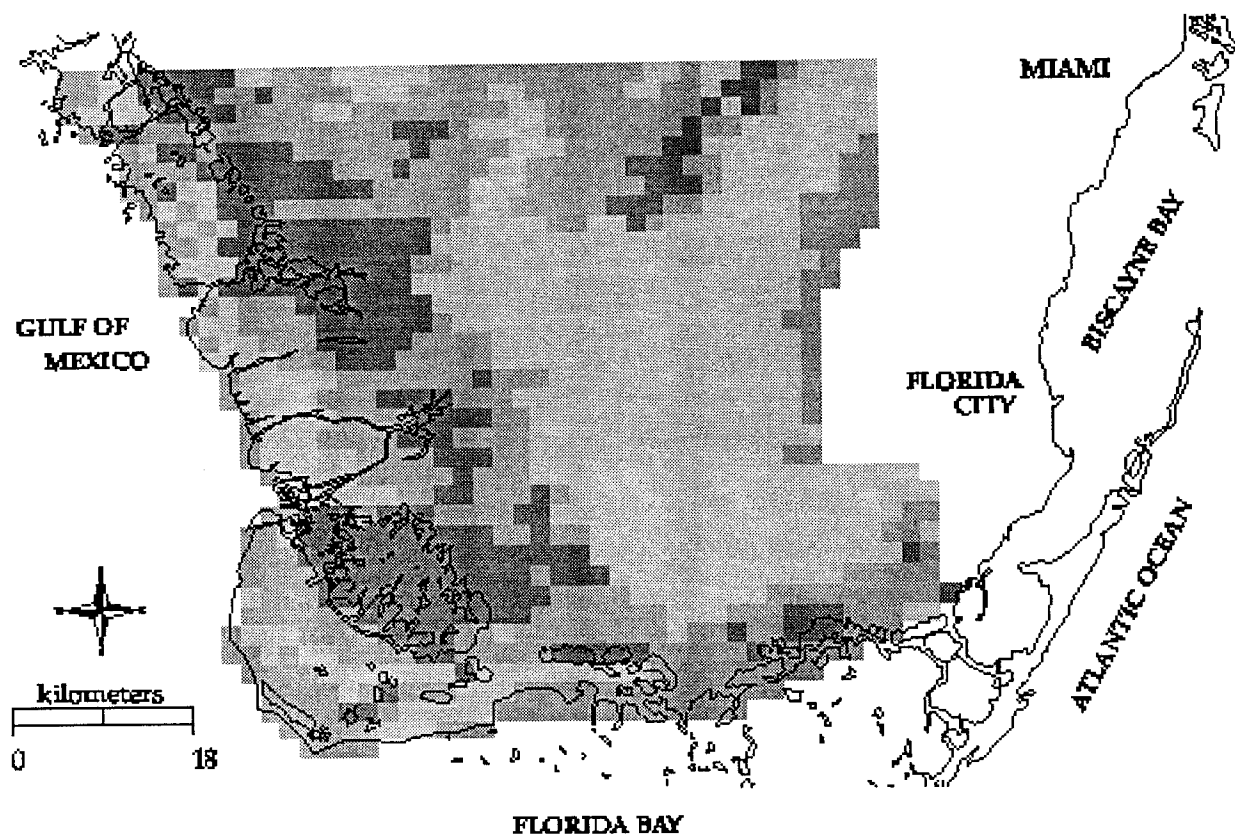
**Observed Class:**

- Incomplete Temporal Record,
Not In Estimate

Surface Water Condition

- Dry
 ▨ Transitional
 ■ Wet

DRAFT COPY



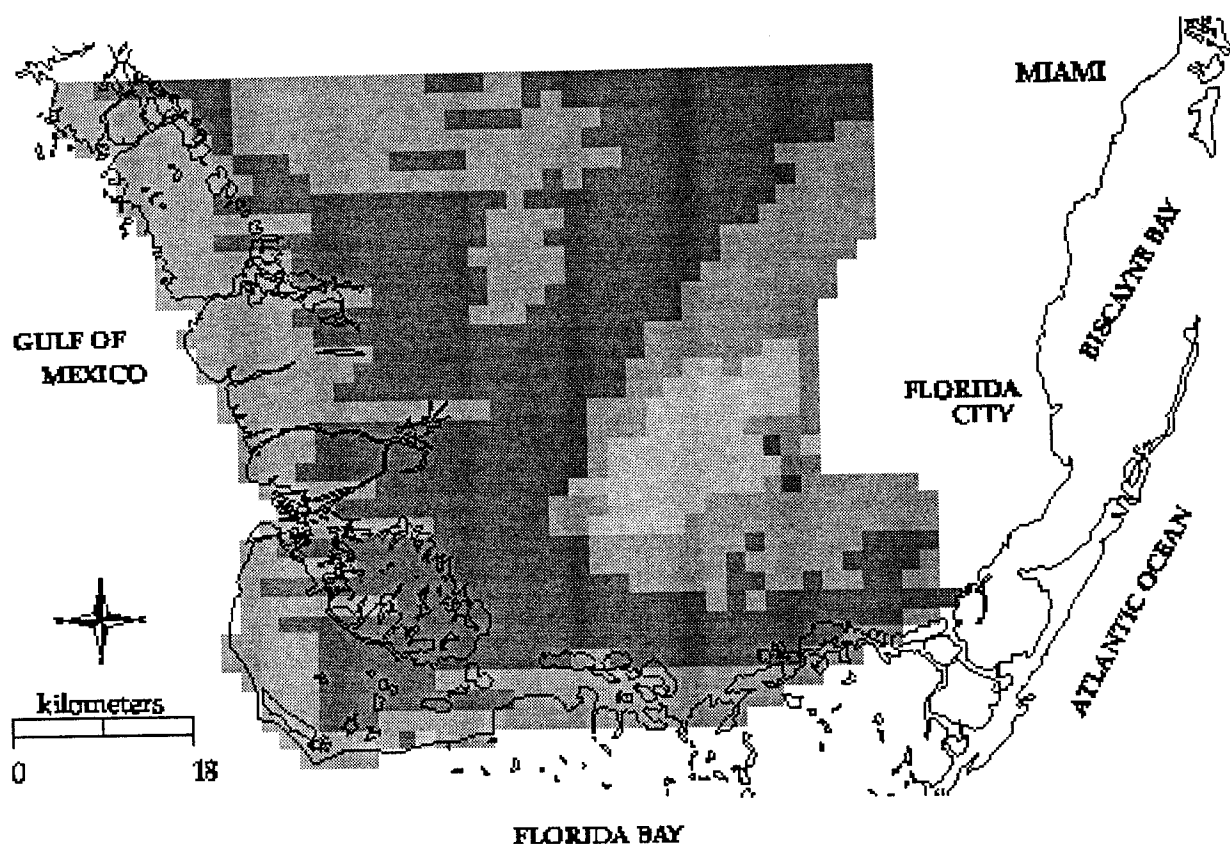
Observed Class:

■ Incomplete Temporal Record,
Not In Estimate

Surface Water Condition

□ Dry
▨ Transitional
■ Wet

DRAFT COPY



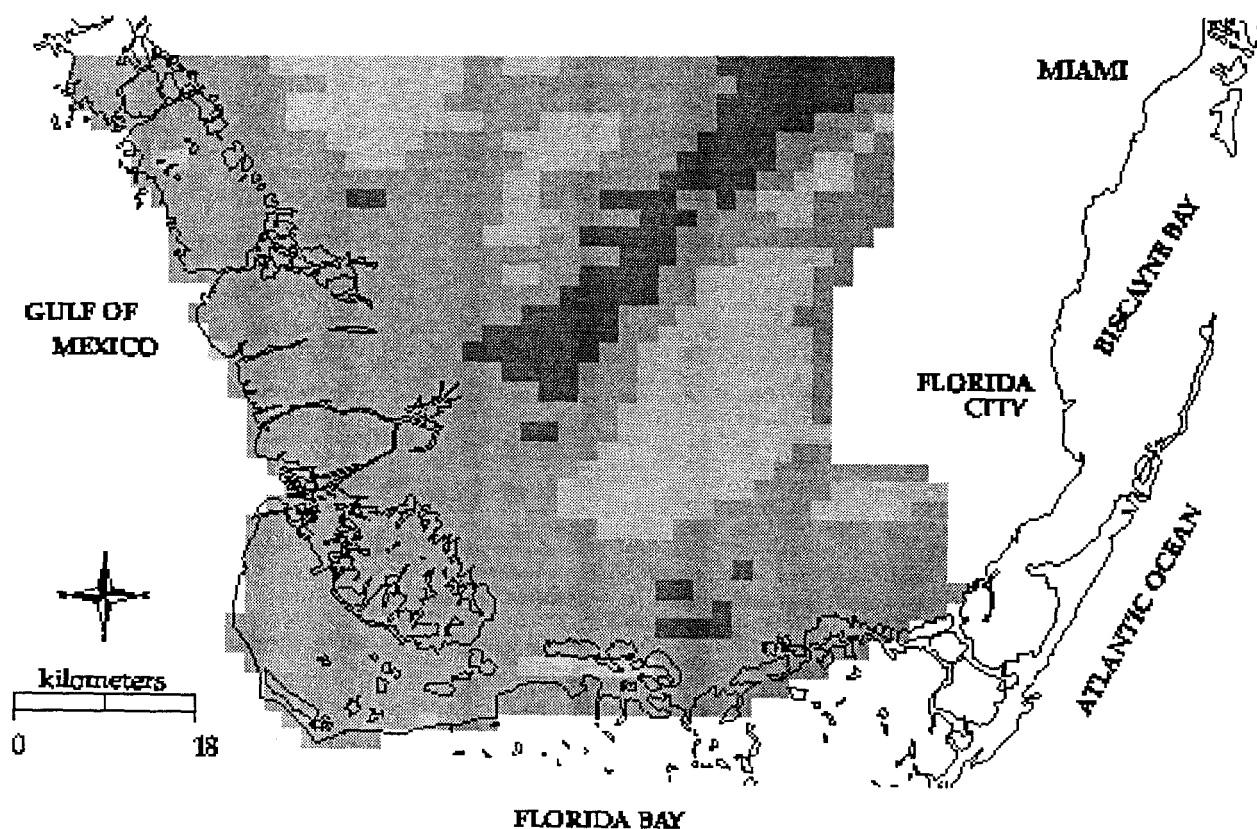
Observed Class:

■ Incomplete Temporal Record,
Not In Estimate

Surface Water Condition

■ Dry
■ Transitional
■ Wet

DRAFT COPY



Observed Class:

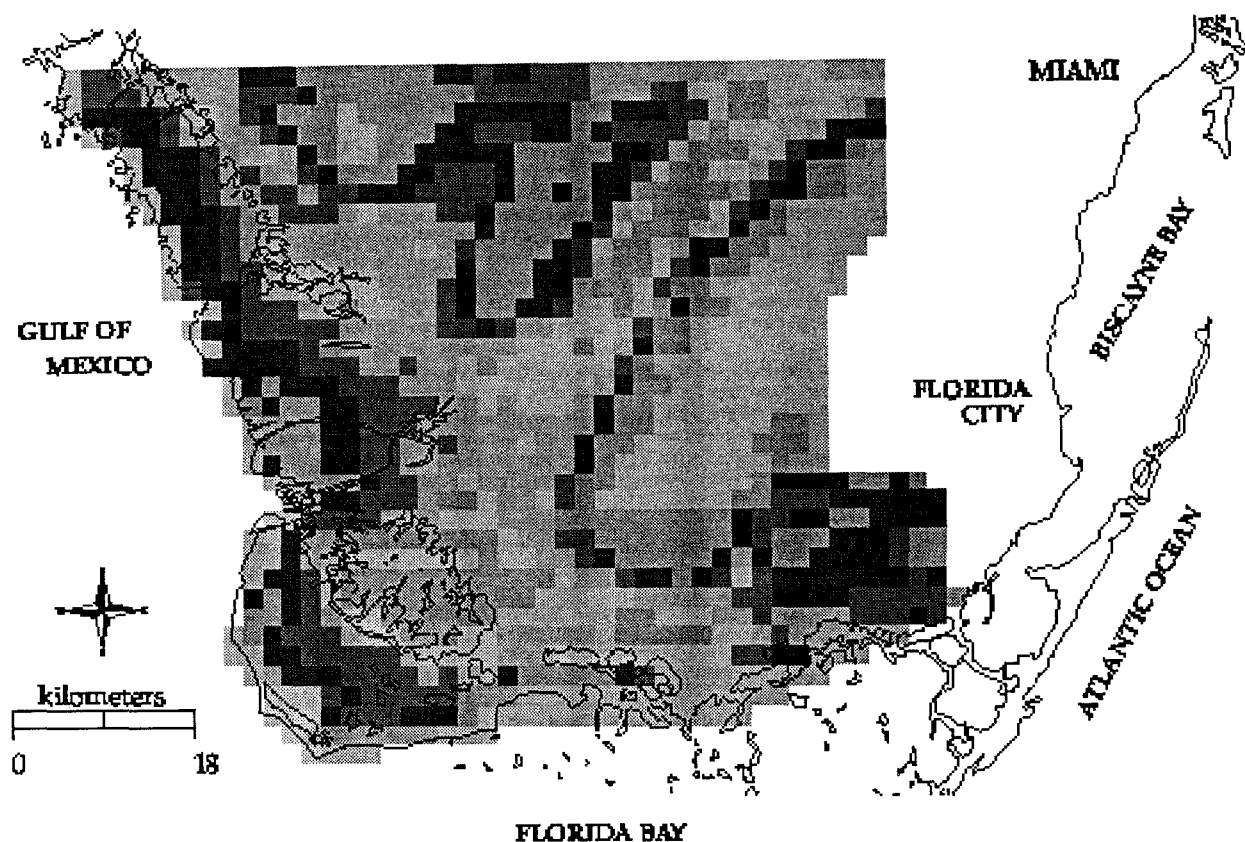
- Incomplete Temporal Record,
Not In Estimate

Surface Water Condition

- Dry
- Transitional
- Wet

DRAFT COPY

17.15 Figure 15. Index of water level condition, SRF study area



Observed Class: Incomplete Temporal Record, Not In Estimate

Surface Water Condition Index

 0	 $46 \leq X \leq 60$
 $1 \leq X \leq 15$	 $61 \leq X \leq 75$
 $16 \leq X \leq 30$	 $76 \leq X \leq 90$
 $31 \leq X \leq 45$	

DRAFT COPY

- 19.16 Figure 16: Plots of wading bird observations for four time points representing four very different spatial patterns
- 19.17 Figure 17: Time series plot of all species estimated abundance and 95% confidence bars
- 19.18 Figure 18: Time series plot of long legged tactile feeders estimated abundance and 95% confidence bars. Time series plots of individual species in this guild are also provided
- 19.19 Figure 19: Time series plot of long legged visual feeders estimated abundance and 95% confidence bars. Time series plots of individual species in this guild are also provided
- 19.20 Figure 20: Time series plot of short legged visual feeders estimated abundance and 95% confidence bars. Time series plots of individual species in this guild are also provided
- 19.21 Figure 21: Time series plot of small dark and small white heron estimated abundance and 95% confidence bars
- 19.22 Figure 22: Water condition fractions pattern as a simplex
- 19.23 Figure 23: Weighted kappa statistics on water condition associations
- 19.24 Figure 24: Estimated extradispersion by date for all species combined
- 19.25 Figure 25: Estimated extradispersion by date for long legged tactile feeders
- 19.26 Figure 26: Estimated extradispersion by date for long legged visual feeders
- 19.27 Figure 27: Estimated extradispersion by date for short legged visual feeders
- 19.28 Figure 28: Estimated extradispersion by date for small dark and small white herons
- 19.29 Figure 29: Map of the Everglades area, showing the boundaries of protected areas

[I will make this one!]

DRAFT COPY

20 Tables

20.1 Table 1: Estimated percent observed by species for SRF pilot study

Species	Expected Percent Detected
All species	
Long-legged Tactile Feeders	
Wood Stork	
Roseate Spoonbill	
Long-legged visual feeders	
Great Blue Herons	
Great White Herons	
Great Egrets	
Short-legged Tactile Feeders	
Glossy Ibis	
White Ibis	
Small Dark Herons	
Small White Herons	

[Do we have these data?]

DRAFT COPY

20.2 Table 2: Land element themes defined from existing maps

Land element Theme	Number of categories	Description
Elevation (Figure 3)	7	Dominant elevation, to a 1ft elevational category (0–7ft msl), for each sample cell determined from contour maps.
Soil type (Figure 4)	7	Dominant soil type, from USDA/SCS soil survey maps for each sample cell determined by the type with maximal areal coverage.
Specific vegetation (Figure 5)	15	Dominant vegetation, based on a structural classification system (open/dense, short/tall, prairie/forest, etc.) developed for the Everglades, for each cell was classified to type with maximal relative areal coverage.
Basin (Figure 6)	13	Hydrological Basins as designated by Everglades National Park and Big Cypress National Preserve/
Physiographic region (Figure 7)	3	Physiographic category for a cell based on the three class system: Mangrove/Estuary, Freshwater Grasslands, Cypress Forests.
Watershed category (Figure 8)	4	Watershed category for a cell based on a four region system: Big Cypress, Eastern Panhandle, Shark Slough, Taylor Slough.

20.3 Table 3: Land element categorizations based on aerial observations

Land element Theme	Number of categories	Description
Pond density (Figure 9)	3	Pond density category for each sample grid cell using three levels: No Ponds, Low Density (1–5 ponds/km ²), Moderate Density (6–10 ponds/km ²).
Gross vegetation (Figure 10)	2	Gross vegetation categorization using two levels: Dense Vegetation Cover and Open Vegetation Cover.
Ground surface (Figure 11)	11	Dominant ground surface features for each sample grid cell using an eleven-category classification scheme.

DRAFT COPY

20.4 Table 4: Landscape classification based on multivariate analysis of land element themes

Landscape classification	Number of categories	Description
HLC (Figure 12)	7	A Hierarchical Landscape Classification was developed using GIS software (GRAS—Gross and Ginfer 19??; Astel <i>et al.</i> 1969) by overlaying many layers for elevation, dominant ground surface features, soil type, and dominant vegetation cover type with an index of water level condition, then using system functions to identify boundaries of generally defined landscape types.
SCC (Figure 13)	14	A Statistical CLuster Classification was developed using statistical hierarchical cluster analysis techniques (Johnson and Wichern 1992) using eight of the Land Element Theme categorizations (excluding Gross Vegetation). The method reduced the number of possible multidimensional classifications from $7 \times 7 \times 15 \times 13 \times 3 \times 4 \times 3 \times 11$ to 14 by clustering on multidimensional concordance. Categories are defined in Table 5.

20.5 Table 5: Landscape classification based on multivariate analysis of land element themes

Class	Basin	Elevation (max ft msl)	Ground surface features	Pond density	Soil type	Vegetation	Physio- graphic region	Watershed	HLC	Number of cells
1	Big Cypress Mang. Est.	1	inland flats	no ponds	m/m, tidal	mangrove, prairie	mangrove estuaries	Big Cypress	estuary interior flats/bay complex	86
2	Southern Big Cypress	2 & 6	slough/ outcrop	no ponds	marls, sandy	cypress, pine forest	cypress forest	Big Cypress	outcrop/ ridge complex	73
3	Shark Slough	7	ridge	no ponds	marls, sandy	open sawgrass/ muhlen- bergia prairie	freshwater grassland	Shark Slough	outcrop/ ridge complex	58
4	Northeast Shark Slough	6	slough	no ponds	m/m, freshwater	dense sawgrass prairie	freshwater grassland	Shark Slough	central slough	48
5	East Slough	1	slough	no ponds	marls, sandy	tree island, open prairie	freshwater grassland	Shark Slough	high slough/ swamp complex	64
6	Shark Slough	2	slough	no ponds	muck, freshwater	dense sawgrass prairie	freshwater grassland	Shark Slough	central slough	105
7	Northeast Shark Slough/ Shark Slough	6	lowland	no ponds	rock, land	dense sawgrass prairie/ tree island, open prairie	freshwater grassland	Shark Slough	rock land	48
8	Shark Slough Mang. Est.	1	inland flats	no ponds	m/m, total	mangrove, prairie	mangrove estuaries	Shark Slough	estuary interior flats/bay complex	329
9	Northeast Taylor Slough	6	lowland	no ponds	rock, land	tree island, open prairie	freshwater grassland	Taylor Slough	rock land	58
10	LPK/S taylor/ Geo Area	2	ridge	no ponds	marls, organic/ calcium	tree island, open prairie	freshwater grassland	Taylor Slough	outcrop/ ridge complex	106
11	Eastern Panhandle	2	lowland	no ponds	marls, organic/ calcium	open sawgrass/ muhlen- bergia prairie	freshwater grassland	Eastern Panhandle	rock land	33
12	LPK/STS/ EP/Mang. Est.	1	inland flats	no ponds	m/m, tidal	tall/short mangrove forest	mangrove estuaries	Taylor Slough	outcrop/ ridge complex	50
13	Eastern Panhandle/ Mang. Est.	1	coastal tidal flats	no ponds	m/m, tidal	short mangrove forest	mangrove estuaries	Eastern Panhandle	outcrop/ ridge complex	16
14	Northeast Taylor Slough	3	slough	low	rock, land	dense sawgrass prairie	freshwater grassland	Taylor Slough	high slough/ swamp complex	1

DRAFT COPY

20.6 Table 6: Models used to examine various dependence structures with the bird count data

Model Name	Description
Independence	Bird counts between any two transect/cells are assumed independent. This is equivalent to calculating the sample mean count and standard deviation assuming all 1075 cells are independent observations from a Poisson distribution.
Compound symmetry	Bird counts between any two cells located on the same transect are assumed correlated. Counts between cells in different transect are assumed not correlated. The model assumes that cells along the same transect are 'interchangeable' with each other as far as counts are concerned. This model assumes a dependency structure, but it is not very strong.
Autoregressive AR(1)	Counts of cells located next to each other along a transect exhibit a higher degree of dependency than do cells located further apart. Thus, if counts from two neighboring cells have a correlation of $-1 < \rho < 1$, cells that are separated by a cell will have count correlation of ρ^2 , cells separated by two cells will have count correlations of ρ^3 , etc. Cells in adjacent transects are assumed uncorrelated.
Autoregressive AR(1) with random transects	The same correlation structure as for AR(1) with the addition that transect totals are considered random. This has the property that counts from cells in adjacent transect are assumed correlated.
Spherical, no nugget effect	Cell counts along and between transects are assumed correlated, with the level of correlation decreasing the further the cells are apart. In geostatistics this model is referred to as the isotropic correlation model since the correlation between two cells does not depend on compass direction, only distance.
Spherical, with nugget effect	Cell observation correlations follow the same structure as the 'no-nugget' model, but with a significant degree of non-zero background variability (the nugget effect). This has the effect of assuming that the variance of counts is higher than in the no-nugget model.

20.7 Table 7: Schwartz's Bayesian Criterion values for alternate GLMM models for selected dates, all species combined

Date YY-MM	Indepen- dence	Compound symmetry	AR(1)	AR(1), random transect effect	Spherical, no nugget	Spherical, with nugget
85-01	-1758.5	-1724.9	-1631.1	-1631.6	-1618.9	-1569.9
89-04	-920.0	-915.1	-847.7	-845.2	-839.0	-828.2
92-12	-1655.3	-1632.5	-1599.1	-1596.1	-1591.3	-1564.5
95-03	-1459.0	-1447.0	-1393.4	-1396.2	-1395.1	-1356.1

DRAFT COPY

20.8 Table 8: Total counts and percentages by feeding guild and species

Group	Count	Percent
All Species	495086	100.00
Long Legged Tactile Feeders	25875	5.23
Wood Stork	20681	4.18
Roseate Spoonbill	5194	1.05
Long Legged Visual Feeders	167189	33.77
Great Blue Herons	9778	1.98
Great White Herons	2335	.47
Great Egrets	155076	31.32
Short Legged Tactile Feeders	247652	50.02
Glossy Ibis	5883	1.19
White Ibis	241769	48.83
Small Dark Herons	21410	4.32
Small White Herons	32960	6.66

DRAFT COPY

20.9 Table 9: Minimum and maximum estimated species abundance by feeding guild and species

Group	Estimated abundance			
	Minimum	Date	Maximum	Date
All Species	1327	5/90	90035	1/93
Long Legged Tactile Feeders	7	8/86	10401	1/94
Wood Stork	7	8/95	9825	1/94
Roseate Spoonbill	7	8/86	3103	4/92
Long Legged Visual Feeders	352	5/90	26484	1/94
Great Blue Herons	33	5/90	2444	3/85
Great White Herons	7	5/90	1304	2/89
Great Egrets	313	5/90	25644	1/94
Short Legged Tactile Feeders	513	4/95	50726	2/85
Glossy Ibis	7	7/85	3527	2/85
White Ibis	500	4/95	49562	1/93
Small Dark Herons	27	4/90	5002	3/94
Small White Herons	67	11/85	7858	1/92

DRAFT COPY

20.10 Table 10: Jaccard index for association, averaged over time, between guilds, with associated variance of this index over time measured as the standard deviation

Guilds		Average Jaccard Index (Standard Deviation)
Long-legged Tactile	Long-legged Visual	0.094 (0.077)
Long-legged Tactile	Short-legged Tactile	0.108 (0.079)
Long-legged Visual	Short-legged Tactile	0.236 (0.123)

DRAFT COPY

20.11 Table 11: Jaccard index for association, averaged over time, between species, with associated variance of this index over time measured as the standard deviation

Species		Average Jaccard Index (Standard Deviation)
gbh	ge	ne
gbh	gi	0.019 (0.015)
gbh	gwh	0.024 (0.026)
gbh	rs	0.023 (0.027)
gbh	sdh	0.101 (0.060)
gbh	swh	0.098 (0.063)
gbh	wi	0.235 (0.124)
gbh	ws	0.095 (0.075)
ge	gi	ne
ge	gwh	0.003 (0.010)
ge	rs	0.017 (0.067)
ge	sdh	0.037 (0.033)
ge	swh	0.015 (0.021)
ge	wi	0.035 (0.026)
ge	ws	0.020 (0.026)
gi	gwh	ne
gi	rs	0.064 (0.012)
gi	sdh	0.021 (0.027)
gi	swh	0.023 (0.032)
gi	wi	0.028 (0.033)
gi	ws	0.020 (0.035)
gwh	rs	ne
gwh	sdh	0.030 (0.027)
gwh	swh	0.034 (0.036)
gwh	wi	0.033 (0.029)
gwh	ws	0.045 (0.061)

DRAFT COPY

rs	sdh	ne
rs	swh	0.125 (0.068)
rs	wi	0.135 (0.060)
rs	ws	0.060 (0.048)
sdh	swh	ne
sdh	wi	0.104 (0.053)
sdh	ws	0.062 (0.043)
swh	wi	ne
swh	ws	0.099 (0.075)
wi	ws	ne

DRAFT COPY

20.12 Table 12: Decomposition of total variability in bird counts

Group	Total Variation	Model Factor (% of total variability)			
		Year Trend		Cycle W/I year	Residual (spatial and unknown)
		Linear	Quadratic		
All Species	2 223 344	0.01	0.90	4.88	94.21
Long Legged Tactile Feeders	221 646	0.10	0.06	5.31	94.53
Wood Stork	na*	na	na	na	na
Roseate Spoonbill	na	na	na	na	na
Long Legged Visual Feeders	699 473	0.004	0.01	3.66	96.33
Great Blue Herons	57 904	0.47	0.78	4.15	94.60
Great White Herons	na	na	na	na	na
Great Egrets	683 051	0.02	0.01	3.46	96.51
Short Legged Tactile Feeders	1 553 100	0.08	0.31	5.15	94.46
Glossy Ibis	na	na	na	na	na
White Ibis	1 519 668	0.08	0.32	5.06	94.54
Small Dark Herons	160 763	0.51	0.004	1.12	98.37
Small White Herons	266 720	0.30	0.004	0.56	99.14

*na = not available, model failed to converge. This is usually due to sparse data distributed across all dates, or data concentrated in just a few months.

DRAFT COPY

20.13 Table 13: Importance of land element themes in describing spatial pattern

Group	Land Element Theme (% of total variability)							
	Basin	Elevation	Soil Type	Vegetation	Physio-graphic Region	Pond Density	Ground Surface Features	Watershed
All Species	6.34	0.73	2.24	3.23	0.06	0.20	3.49	2.50
Long Legged Tactile Feeders	5.65	2.28	1.55	3.73	1.59	0.06	3.83	0.42
Wood Stork	na*	na	na	na	na	na	na	na
Roseate Spoonbill	na*	na	na	na	na	na	na	na
Long Legged Visual Feeders	4.67	0.64	1.38	3.53	0.21	0.10	3.12	1.86
Great Blue Herons	5.67	0.48	3.84	5.65	1.06	0.40	4.10	2.20
Great White Herons	na*	na	na	na	na	na	na	na
Great Egrets	4.33	0.65	1.21	3.30	0.19	0.10	2.90	1.80
Short Legged Tactile Feeders	6.26	0.58	2.55	3.00	0.03	0.35	3.65	2.48
Glossy Ibis	na*	na	na	na	na	na	na	na
White Ibis	6.11	0.61	2.43	2.90	0.01	.031	3.59	2.51
Small Dark Herons	6.01	0.65	2.68	2.91	0.22	0.41	2.21	2.92
Small White Herons	6.23	2.48	2.64	3.50	1.82	0.21	3.41	2.26

*na = not available, model failed to converge. This is usually due to sparse data distributed across all dates, or data concentrated in just a few months.

20.14 *Table 14: Patterns of significant over- and under-utilization of cells
with different land element themes*

This table has not been computed at this time. We (KMP and...?) are still not certain how to present this material since it would require a one-page table (at least) for each land element theme.

DRAFT COPY

20.15 Table 15: Importance of HLC and SCC in describing spatial patterns

Group	% of Total Variation	
	Hierarchical Landscape Classification	Statistical Cluster Classification
All Species	3.51	5.91
Long Legged Tactile Feeders	3.18	4.57
Wood Stork	na*	na
Roseate Spoonbill	na	na
Long Legged Visual Feeders	3.80	4.40
Great Blue Herons	5.96	6.27
Great White Herons	na	na
Great Egrets	3.53	4.15
Short Legged Tactile Feeders	3.30	6.31
Glossy Ibis	na	na
White Ibis	3.16	6.15
Small Dark Herons	2.88	5.16
Small White Herons	2.96	4.32

*na = not available, model failed to converge. This is usually due to sparse data distributed across all dates, or data concentrated in just a few months.

DRAFT COPY

20.16 Table 16: Importance of surface water condition in describing temporal and spatial patterns

Group	Total variation	Surface Water Condition Factor (% of total variability)	
		Alone in model	After year trend, annual cycle and basin effect accounted for (see Tables 11 and 12)
All Species	2 223 344	5.14	3.75
Long Legged Tactile Feeders	221 646	1.94	0.86
Wood Stork	na*	na	na
Roseate Spoonbill	na	na	na
Long Legged Visual Feeders	699 473	4.82	3.15
Great Blue Herons	57 904	2.79	2.19
Great White Herons	na		na
Great Egrets	683 051	4.67	3.02
Short Legged Tactile Feeders	1 553 100	4.47	3.57
Glossy Ibis	na	na	na
White Ibis	1 519 668	4.46	3.54
Small Dark Herons	160 763	3.82	3.31
Small White Herons	266 720	2.42	1.11

*na = not available, model failed to converge. This is usually due to sparse data distributed across all dates, or data concentrated in just a few months.

DRAFT COPY

20.17 Table 17: Patterns of significant over- and under-utilization of cells with surface water classifications

Month	Guild			
	All species	Long Legged Tactile Feeders	Long Legged Visual Feeders	Short Legged Tactile Feeders
Jan	D----- I+-+o++++oo++o W+-++++++oo-o	D-----oo I---oo++--o+- Wo++o+--+o+-	D----- Ioo-oo++o-+- W+o+++++o+-	D----- I++++++o++ W+-o++++-o--
Feb	D----- I+o+o++++o+++ W++++++o--	D--o-----o Io--o+++ooo+o W+o+o+o+o+-	D----- I+-oo+++o-+- W++++++o+-o	D----- I+++o++++o+++ W++++++o+-o
Mar	D-----+--- I++-+++++o+++ W+++++o+--+o	D-----o---o- Ioo-++o-+oo+o W+o+ooo+o+-o	D-----+--- I++-+++++--o Wo+++++o+-+-	D-----+o-+ I+++o++++o+++ W+++++o+o+--
Apr	D----- I+-i++++++ W+-+oo+-----	D-o-----o-o+ I++++++oo W+---oo-----	D-o----- I+-++++++o W++oooo+o-o	D----- I+o++++++o+ W+o-oo+---o-
May	D----- I+++o++++++ Wooo+oo+--++	D----ooo-+-o- I-+ooooo+-o-o W+-o+ooo-+o+	D----- I++oo++++++o Wo+o+ooo-+-+	D----- I+++-++++++ W-oo+o+-oo--
Jun	D-o I+- Wo+	D-o I+o Wo-	D-- I+- Wo+	D-o I+- Wo+
Jul	D-- I++ Wo-	Doo Ioo Wo-	D-- I-- W++	D-- I++ W--
Aug	D----+---o- I+++o-+++++ W-o-oo-----	Dooooooooo Ioooo++++oo Woooo--oo	Do--oo---oo Iooo--oo--o Wooo++oo++o	D---oo---oo I+++o+++++ W---o-----
Sep	D--- Io+- Wo--	Dooo Io-o Wooo	D--- I-o- W+o+	D--- Io+- Wo+
Oct	D--- Io+- Wo--	Dooo Iooo Wooo	D--- I-o+ W+o-	D--- I+++ Wo--
Nov	D-o- Io++ Wo--	Dooo Iooo Wooo	D-o- I-o+ W+oo	D-o- I+++ W---
Dec	D----- I-++++-+++ W++o++++-o-	Do-----oooo I+-o+oo-+o+ Woo+-+o+-o-	D-----oo- I--o+---+++ W++++++-o-	D-----oo I-o++++o+-+ W+-++o-o-

DRAFT COPY